Fundamentals of Bayesian Epistemology 1

Introducing Credences

MICHAEL G. TITELBAUM

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First Edition published in 2022

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2021949533

ISBN 978–0–19–870760–8 (hbk.)
ISBN 978–0–19–870761–5 (pbk.)

DOI: 10.1093/oso/9780198707608.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Contents

VOLUME 2

III. APPLICATIONS

IV. ARGUMENTS FOR BAYESIANISM

4

Updating by Conditionalization

Up to this point we have discussed *synchronic* credence constraints—rationally required relations among the degrees of belief an agent assigns at a given time. This chapter introduces the fifth (and final) core normative Bayesian rule, Conditionalization. Conditionalization is a *diachronic* rule, requiring an agent's degrees of belief to line up in particular ways across times.

I begin by laying out the rule and some of its immediate consequences. We will then practice applying Conditionalization using Bayes's Theorem. Some of Conditionalization's consequences will prompt us to ask what notions of learning and evidence pair most naturally with the rule. I will also explain why it's important to attend to an agent's *total* evidence in evaluating her responses to learning.

Finally, we will see how Conditionalization helps Bayesians distinguish two influences on an agent's opinions: the content of her evidence, and her tendencies to respond to evidence in particular ways. This will lead to Chapter 5's discussion of whether multiple distinct responses to the same evidence might ever be rationally permissible. Differing answers to that question provide a crucial distinction between Subjective and Objective Bayesianism.

4.1 Conditionalization

Suppose I tell you I just rolled a fair six-sided die, and give you no further information about how the roll came out. Presumably you assign equal unconditional credence to each of the six possible outcomes, so your credence that the die came up six will be 1/6. I then ask you to suppose that the roll came up even (while being very clear that this is just a supposition—I'm still not revealing anything about the actual outcome). Applying the Ratio Formula to your unconditional distribution, we find that rationality requires your credence in six conditional on the supposition of even to be 1/3. Finally, I break down and tell you that the roll actually did come up even. Now how confident should you be that it came up six?

Fundamentals of Bayesian Epistemology 1: Introducing Credences. Michael G. Titelbaum, Oxford University Press.
 DOI: 10.1093/oso/9780198707608.003.0004

I hope the obvious answer is 1/3. When you learn that the die actually came up even, the effect on your confidence in a six is identical to the effect of merely supposing evenness. This relationship between learning and supposing is captured in Bayesians' credence-updating rule:

Conditionalization: For any time t_i and later time t_j, if proposition E in $\mathcal{L}$ represents everything the agent learns between t_i and t_j, and $\text{cr}_i(E) > 0$, then for any H in $\mathcal{L}$,

$$\text{cr}_j(H) = \text{cr}_i(H \mid E)$$

where cr_i and cr_j are the agent's credence distributions at the two times. Conditionalization captures the idea that an agent's credence in H at t_j—after *learning* E—should equal her earlier t_i credence in H had she merely *supposed* E. If we label the two times in the die-roll case t_1 and t_2, and let 6 represent the die's coming up six and E represent its coming up even, then Conditionalization tells us

$$\text{cr}_2(6) = \text{cr}_1(6 \mid E) \tag{4.1}$$

which equals 1/3 (given what we know about your unconditional distribution at t_1).

> **Warning**
>
> Some theorists take Conditionalization to *define* conditional credence. For them, to assign the conditional credence $\text{cr}_i(H \mid E) = r$ *just is* to be disposed to assign $\text{cr}_j(H) = r$ should you learn E. As I said in Chapter 3, I take conditional credence to be a genuine mental state, manifested by the agent in various ways at t_i (what she'll say in conversation, what sorts of bets she'll accept, etc.) beyond just her dispositions to update. For us, Conditionalization represents a *normative* constraint relating the agent's unconditional credences at a later time to her conditional credences earlier on.

Combining Conditionalization with the Ratio Formula gives us

$$\text{cr}_j(H) = \text{cr}_i(H \mid E) = \frac{\text{cr}_i(H \,\&\, E)}{\text{cr}_i(E)} \tag{4.2}$$

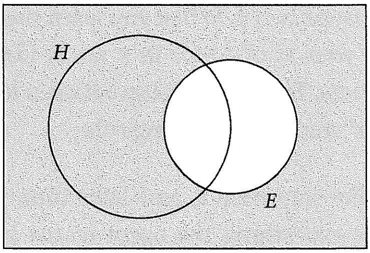

Figure 4.1 Updating on E

(when $\text{cr}_i(E) > 0$). A Venn diagram shows why dividing these particular t_i credences should yield the agent's credence in H at t_j. In Chapter 3 we used a diagram like Figure 4.1 to understand conditional credences. There the white circle represented a set of possibilities to which the agent had temporarily narrowed her focus in order to entertain a supposition.

Now let's imagine the rectangle represents all the possible worlds the agent entertains at t_i (her doxastically possible worlds at that time). The size of the H-circle represents the agent's unconditional t_i credence in H. Between t_i and t_j the agent learns that E is true. Among the worlds she had entertained before, the agent now excludes all the non-E worlds. Her set of doxastic possibilities narrows down to the E-circle; in effect, *the E-circle becomes the agent's new rectangle*. How unconditionally confident is the agent in H now? That depends what fraction of her new doxastic space is occupied by H-worlds. And this is what Equation (4.2) calculates: it tells us what fraction of the E-circle is occupied by H & E worlds.

As stated, the Conditionalization rule is useful for calculating a single unconditional credence value after an agent has gained evidence. But what if you want to generate the agent's entire t_j credence distribution at once? We saw in Chapter 2 that a rational agent's entire t_i credence distribution can be specified by a probability table that gives the agent's unconditional t_i credence in each state-description of $\mathcal{L}$. To satisfy the probability axioms, the credence values in a probability table must be non-negative and sum to 1. The agent's unconditional credence in any (non-contradictory) proposition can then be determined by summing her credences in the state-descriptions on which that proposition is true.

When an agent updates her credence distribution by applying Conditionalization to some learned proposition E, we will say that she "conditionalizes on E". (Some authors say she "conditions on E".) To calculate the probability table resulting from such an update, we apply a two-step process:

1. Assign credence 0 to all state-descriptions inconsistent with the evidence learned.
2. Multiply each remaining nonzero credence by the *same* constant so that they all sum to 1.

As an example, let's consider what happens to your confidence that the fair die roll came up prime[1] when you learn that it came up even:

P	E	cr_1	cr_2
T	T	1/6	1/3
T	F	1/3	0
F	T	1/3	2/3
F	F	1/6	0

Here we've used a language $\mathcal{L}$ with atomic propositions P and E representing "prime" and "even". The cr_1 column represents your unconditional credences at time t_1, while the cr_2 column represents your t_2 credences. Between t_1 and t_2 you learn that the die came up even. That's inconsistent with the second and fourth state-descriptions, so in the first step of our update process their cr_2-values go to 0. The cr_1-values of the first and third state-descriptions (1/6 and 1/3 respectively) add up to only 1/2. So we multiply both of these values by 2 to obtain unconditional t_2-credences summing to 1.[2]

In this manner, we generate your unconditional state-description credences at t_2 from your state-description credences at t_1. We can then calculate cr_2-values for other propositions. For instance, adding up the cr_2-values on the lines that make P true, we find that

$$\mathrm{cr}_2(P) = 1/3 \tag{4.3}$$

Given your initial distribution, your credence that the die came up prime after learning that it came up even is required to be 1/3. Hopefully that squares with your intuitions about what's rationally required in this case!

One final note: Our two-step process for updating probability tables yields a handy fact. Notice that in the second step of the process, every state-description that hasn't been set to zero is multiplied by the *same* constant. When two values are multipled by the same constant, the ratio between them remains intact. This means that if two state-descriptions have nonzero credence values after an update by Conditionalization, those values will stand in the same ratio as they did before the update. This fact will prove useful for problem-solving later on. (Notice that it applies only to *state-descriptions*; propositions that are not state-descriptions may not maintain their credence ratios after a conditionalization.)

4.1.1 Consequences of Conditionalization

If we adopt Conditionalization as our updating norm, what follows? When an agent updates by conditionalizing on E, her new credence distribution is just her earlier distribution conditional on E. In Section 3.1.2 we saw that if an agent's credence distribution obeys the probability axioms and Ratio Formula, then the distribution she assigns conditional on any particular proposition (in which she has nonzero credence) will be probabilistic as well. This yields the important result that if an agent starts off obeying the probability axioms and Ratio Formula and then updates by Conditionalization, her resulting credence distribution will satisfy the probability axioms as well.[3]

The process may then iterate. Having conditionalized her probabilistic distribution cr_1 on some evidence E to obtain probabilistic credence distribution cr_2, the agent may then gain further evidence E', which she conditionalizes upon to obtain cr_3 (and so on). Moreover, conditionalization has the elegant mathematical property of being **cumulative**: Instead of obtaining cr_3 from cr_1 in two steps—first conditionalizing cr_1 on E to obtain cr_2, then conditionalizing cr_2 on E' to obtain cr_3—we can generate the same cr_3 distribution by conditionalizing cr_1 on $E \& E'$, a conjunction representing all the propositions learned between t_1 and t_3. (You'll prove this in Exercise 4.3.) Because conditionalization is cumulative it is also **commutative**: Conditionalizing first on E and then E' has the same effect as conditionalizing in the opposite order.

Besides being mathematically elegant, cumulativity and commutativity are intuitively plausible features of a learning process. Suppose a detective investigating a crime learns that the perpetrator was an Italian accordionist, and updates her credences accordingly. Intuitively, it shouldn't matter if we describe this episode as the detective's learning first one piece of evidence and then another (first that the perpetrator was Italian, and then that he was an accordionist) or as the detective's learning a single conjunction containing both. Because conditionalization is cumulative, it will prescribe the same ultimate credences for the detective on either construal. Similarly, it shouldn't matter whether we take her to have learned that the perpetrator was an Italian accordionist or an accordion-playing Italian. Because conditionalization is commutative, the order in which pieces of evidence are presented makes no difference to an agent's ultimate credences.[4]

When an agent conditionalizes on evidence E, what happens to her unconditional credence in that very evidence? Substituting E for H in Equation (4.2) (and recalling that $E \& E$ is equivalent to E), we can see that if an agent learns E between t_i and t_j then

$$\mathrm{cr}_j(E) = 1 \tag{4.4}$$

Conditionalization creates certainties; conditionalizing on a piece of evidence makes an agent certain of that evidence. Moreover, any proposition entailed by that evidence must receive at least as high a credence as the evidence (by our Entailment rule). So an agent who conditionalizes also becomes certain of any proposition entailed by the evidence she learns.

And conditionalization doesn't just create certainties; it also retains them. If an agent is certain of a proposition at t_i and updates by Conditionalization, she will remain certain of that proposition at t_j. That is, if $\mathrm{cr}_i(H) = 1$ then Conditionalization yields $\mathrm{cr}_j(H) = 1$ as well. On a probability table, this means that once a state-description receives credence 0 at a particular time (the agent has ruled out that possible state of the world), it will receive credence 0 at all subsequent times as well.

In Exercise 4.2 you'll prove that conditionalizing retains certainties from the probability axioms and Ratio Formula. But it's easy to see why this occurs on a Venn diagram. You're certain of H at t_i when H is true in every world you consider a live doxastic possibility. Conditionalizing on E strictly narrows the set of possible worlds you entertain. So if H was true in every world you entertained before conditionalizing, it'll be true in every world you entertain afterwards as well.

Combining these consequences of Conditionalization yields a somewhat counterintuitive result, to which we'll return in later discussions. Conditionalizing on E between two times makes that proposition (and any proposition it entails) certain. Future updates by Conditionalization will then retain that certainty. So if an agent updates by conditionalizing throughout her life, any piece of evidence she learns at any point will remain certain for her ever after.

What if an agent doesn't learn *anything* between two times? Bayesians represent an empty evidence set as a tautology. So when an agent gains no information between t_i and t_j, Conditionalization yields

$$\mathrm{cr}_j(H) = \mathrm{cr}_i(H \mid \top) = \mathrm{cr}_i(H) \tag{4.5}$$

for any H in $\mathcal{L}$. (The latter half of this equation comes from Equation (3.7), which noted that credences conditional on a tautology equal unconditional credences.) If an agent learns nothing between two times and updates by Conditionalization, her degrees of confidence remain unchanged.

4.1.2 Probabilities are weird! The Base Rate Fallacy

Bayes's Theorem expresses a purely synchronic relation; as we saw in Section 3.1.3, for any time t_i it calculates $\mathrm{cr}_i(H \mid E)$ in terms of other credences assigned at that time. Nevertheless, our diachronic Conditionalization rule gives Bayes's Theorem added significance. Conditionalization says that your unconditional t_j credence in hypothesis H after learning E should equal $\mathrm{cr}_i(H \mid E)$. Bayes's Theorem is a tool for calculating this crucial value (your "posterior" at t_i) from other credences you assign at t_i. As new evidence comes in over time and we repeatedly update by conditionalizing, Bayes's Theorem can be a handy tool for generating new credences from old.

For example, we could've used Bayes's Theorem to answer our earlier question of what happens to your credence in six when you learn that a fair die roll has come up even. The hypothesis is 6, and the evidence is E (for even). By Conditionalization and then Bayes's Theorem,

$$\mathrm{cr}_2(6) = \mathrm{cr}_1(6 \mid E) = \frac{\mathrm{cr}_1(E \mid 6) \cdot \mathrm{cr}_1(6)}{\mathrm{cr}_1(E)} \tag{4.6}$$

$\mathrm{cr}_1(6)$, your prior credence in a six, is $1/6$, and $\mathrm{cr}_1(E)$, your prior credence in even, is $1/2$. The likelihood of E, $\mathrm{cr}_1(E \mid 6)$, is easy—it's 1. So the numerator is $1/6$, the denominator is $1/2$, and the posterior $\mathrm{cr}_1(6 \mid E) = \mathrm{cr}_2(6) = 1/3$ as we saw before.[5]

Let's apply Bayes's Theorem to a more interesting case:

> One in 1,000 people have a particular disease. You have a test for the presence of the disease that is 90% accurate, in the following sense: If you apply the test to a subject who has the disease it will yield a positive result 90% of the time, and if you apply the test to a subject who lacks the disease it will yield a negative result 90% of the time.
>
> You randomly select a person and apply the test. The test yields a positive result. How confident should you be that this subject actually has the disease?

Most people—including trained medical professionals!—answer this question with a value around 80% or 90%. But if you set your credences by the statistics given in the problem, the rationally required degree of confidence that the subject has the disease is less than 1%.

We'll use Bayes's Theorem to work that out. Let D be the proposition that the subject has the disease and P the proposition that when applied to the subject, the test yields a positive result. Here D is our hypothesis, and P is the evidence acquired between t_1 and t_2. At t_1 (before applying the test) we take the subject to be representative of the population, giving us priors for the hypothesis and the catchall:

$$\mathrm{cr}_1(D) = 0.001 \qquad\qquad \mathrm{cr}_1(\sim D) = 0.999$$

The accuracy profile of the test gives us likelihoods for the hypothesis and catchall:

$$\mathrm{cr}_1(P \mid D) = 0.9 \qquad\qquad \mathrm{cr}_1(P \mid \sim D) = 0.1$$

In words, you're 90% confident that the test will yield a positive result given that the subject has the disease, and 10% confident that we'll get a "false positive" on the supposition that the subject lacks the disease.

Now we'll apply a version of Bayes's Theorem from Section 3.1.3, in which the Law of Total Probability has been used to expand the denominator:

$$\begin{aligned}\mathrm{cr}_2(D) &= \frac{\mathrm{cr}_1(P \mid D) \cdot \mathrm{cr}_1(D)}{\mathrm{cr}_1(P \mid D) \cdot \mathrm{cr}_1(D) + \mathrm{cr}_1(P \mid \sim D) \cdot \mathrm{cr}_1(\sim D)} \\ &= \frac{0.9 \cdot 0.001}{0.9 \cdot 0.001 + 0.1 \cdot 0.999} \\ &\approx 0.009 = 0.9\%\end{aligned} \tag{4.7}$$

So there's the calculation. After learning of the positive test result, your credence that the subject has the disease should be a little bit less than 1%. But even having seen this calculation, most people find it hard to believe. Why shouldn't we be more confident that the subject has the disease? Wasn't the test 90% accurate?

Tversky and Kahneman (1974) suggested that in cases like this one, people's intuitive responses ignore the "base rate" of a phenomenon. The base rate in our example is the prior credence that the subject has the disease. In this case, that base rate is extremely low (1 in 1,000). But people tend to forget about the base rate and be overwhelmed by accuracy statistics (such as likelihoods) about the test. This is known as the **Base Rate Fallacy**.

Why is the base rate so important? To illustrate, let's suppose you applied this test to 10,000 people. Using the base rate statistics, we would expect about ten of those people to have the disease. Since the test gives a positive result for 90% of people who have the disease, we would expect these ten diseased people to yield about nine positive results—so-called "true positives". Then there would be about 9,990 people lacking the disease. Since $\text{cr}_i(P \mid {\sim}D)$—the false positive rate—is 10%, we'd expect to get about 999 false positive results. Out of 1,008 positive results the test would yield, only nine of those subjects (or about 0.9%) would actually have the disease. This particular disease is so rare—its base rate is so tiny—that even with an accurate test we should expect the false positives to swamp the true positives. So when a single randomly selected individual takes the test and gets a positive result, we should be much more confident that this is a false positive than a true one.

Another way to see what's going on is to consider the **Bayes factor** of the evidence you receive in this case. Using Conditionalization and the Ratio Formula, we can derive

$$\frac{\text{cr}_j(H)}{\text{cr}_j({\sim}H)} = \frac{\text{cr}_i(H \mid E)}{\text{cr}_i({\sim}H \mid E)} = \frac{\text{cr}_i(H)}{\text{cr}_i({\sim}H)} \cdot \frac{\text{cr}_i(E \mid H)}{\text{cr}_i(E \mid {\sim}H)} \tag{4.8}$$

That last fraction on the right—the ratio of the likelihood of the hypothesis to the likelihood of the catchall—is the Bayes factor. Personally, I found this equation fairly impenetrable until I remembered that $\text{cr}(H)/\text{cr}({\sim}H)$ is an agent's odds for the proposition H (Section 2.3.4). That means we can rewrite Equation (4.8) as

$$\text{odds for } H \text{ after update} = \text{odds for } H \text{ before update} \cdot \text{Bayes factor} \tag{4.9}$$

If you update by Conditionalization, learning E multiplies your odds for H by the Bayes factor. The Bayes factor thus provides a handy way to measure how much learning E affects your opinion about the hypothesis.

In our disease example, the Bayes factor is

$$\frac{\text{cr}_1(P \mid D)}{\text{cr}_1(P \mid {\sim}D)} = \frac{0.9}{0.1} = 9 \tag{4.10}$$

At t_1, your odds for D are $1:999$. Applying the test has a substantial influence on these odds; as the Bayes factor reveals, a positive test result multiplies the odds by 9. This reflects the high accuracy of the test. Yet since the odds were so small initially, multiplying them by 9 only brings them up to $9:999$. So even after seeing the test outcome, you should be much more confident that the subject doesn't have the disease than you are that she does.[6]

4.2 Evidence and certainty

Combining Conditionalization with the probability axioms and Ratio Formula creates a Bayesian approach to evidence that many have found troubling. Conditionalization works with a proposition E representing everything the agent learns between two times. (If many propositions are learned, E is their conjunction.) We also speak of E as the evidence the agent gains between those two times. Yet Conditionalization gives E properties that epistemologists don't typically attribute to evidence.

We've already seen that a piece of evidence E (along with anything it entails) becomes certain once conditionalized upon. When an agent learns E, the set of doxastically possible worlds she entertains shrinks to a set of worlds that all make E true; on the Venn diagram, what once was merely an E-circle *within* her rectangle of worlds now becomes the entire rectangle. And as we saw in Section 4.1.1, this change is permanent: as long as the agent keeps updating by Conditionalization, any evidence she once learned remains certain and possible worlds inconsistent with it remain ruled out.

Is there any realistic conception of evidence—and of learning—that satisfies these conditions? When I learn that my sister is coming over for Thanksgiving dinner, I become highly confident in that proposition. But do I become 100% certain? Do I *rule out* all possible worlds in which she doesn't show, refusing to consider them ever after? As Richard C. Jeffrey put it:

> Certainty is quite demanding. It rules out not only the far-fetched uncertainties associated with philosophical skepticism, but also the familiar uncertainties that affect real empirical inquiry in science and everyday life.
>
> (2004, p. 53)

This concern about certainties motivates the

Regularity Principle: In a rational credence distribution, no logically contingent proposition receives unconditional credence 0.

The Regularity Principle captures the common-sense idea that one's evidence is never so strong as to entirely rule out any logical possibility. (Recall that a logically contingent proposition is neither a logical contradiction nor a logical tautology.)[7] As damning evidence against a contingent proposition mounts up, we may keep decreasing and decreasing our credence in it, but our unconditional credence distribution should always remain **regular**—it should assign each logically contingent proposition at least a tiny bit of confidence.[8]

The Regularity Principle adds to the synchronic Bayesian rules we have seen so far—it is not entailed by the probability axioms, the Ratio Formula, or any combination of them. As our Contradiction result showed in Section 2.2.1, those rules do entail that all logical contradictions receive credence 0. But Regularity is the converse of Contradiction; instead of saying that *all* contradictions receive credence 0, it entails that *only* contradictions do. Similarly, Regularity (along with the probability axioms) entails the converse of Normality: instead of saying that *all* tautologies receive credence 1, it entails that *only* tautologies do. (The negation of a contingent proposition is contingent; if we were to assign a contingent proposition credence 1 its negation would receive credence 0, in violation of Regularity.) This captures the commonsense idea that one should never be absolutely certain of a proposition that's not logically true.[9]

Conditionalization conflicts with Regularity; the moment an agent conditionalizes on contingent evidence, she assigns credence 1 to a non-tautology. As we saw earlier, conditionalizing on contingent evidence rules out doxastic possibilities the agent had previously entertained; on the Venn diagram, it narrows the set of worlds under consideration. Regularity, on the other hand, fixes an agent's doxastic possibility set as the full set of logical possibilities. While evidence might shift the agent's credences around among various possible worlds, an agent who satisfies Regularity will never eliminate a possible world outright.

We might defend Conditionalization by claiming that whenever agents receive contingent evidence, it is of a highly specific kind, and Regularity is false for this kind of evidence. Perhaps I don't actually learn that my sister is coming over for Thanksgiving—I learn that she *told* me she's coming; or that it *seemed* to me that she said that; or that I had a phenomenal experience as of.... Surely I can be certain what phenomenal experiences I've had, or at least what experiences I'm having right now. While in the midst of having a particular phenomenal experience, can't I entirely rule out the logical possibility that I am having a different experience instead? C.I. Lewis defended this approach as follows:

> If anything is to be probable, then something must be certain. The data which themselves support a genuine probability, must themselves be certainties. We do have such absolute certainties, in the sense data initiating belief and in those passages of experience which later may confirm it. (1946, p. 186)

Yet foundationalist epistemologies based on sense data and indubitable phenomenology have become unpopular in recent years. So it's worth considering other ways to make sense of Conditionalization's conception of evidence. Levi (1980) took credence-1 propositions to represent "standards of serious possibility":

> When witnessing the toss of a coin, [an agent] will normally envisage as possibly true the hypothesis that the coin will land heads up and that it will land tails up. He may also envisage other possibilities—e.g., its landing on its edge. However, if he takes for granted even the crudest folklore of modern physics, he will rule out as impossible the coin's moving upward to outer space in the direction of Alpha Centauri. He will also rule out the hypothesis that the Earth will explode. (p. 3)

However, Levi formalized his standards of serious possibility so that they could change—growing either stronger or weaker—for a given agent over time. So his approach did not fully embrace Conditionalization.

Alternatively, we could represent agents as ruling out contingent possibilities only relative to a particular inquiry. Consider a scientist who has just received a batch of experimental data and wants to weigh its import for a set of hypotheses. There are always outlandish possibilities to consider: the data might have been faked; the laws of physics might have changed a moment ago; she might be a brain in a vat. But to focus on the problem at hand, she might conditionalize on the data and see where that takes her credences in the hypotheses. Updating by Conditionalization might fail as a big-picture, permanent strategy, but nevertheless could be useful in carefully delimited contexts. (I mentioned this possibility in Section 2.2.3.)

Perhaps these interpretations of evidence conditionalized-upon remain unsatisfying. We will return to this problem in Chapter 5, and consider a generalized updating rule (Jeffrey Conditionalization) that allows agents to redistribute their credences over contingent possibilities without eliminating any of them entirely. For the rest of this chapter we will simply assume that Conditionalization on some kind of contingent evidence is a rational updating rule, so as to draw out further features of such updates.

4.2.1 Probabilities are weird! Total Evidence and the Monty Hall Problem

Classical entailment is **monotonic** in the following sense: If a piece of evidence E entails H, any augmentation of that evidence (any conjunction that includes E as a conjunct) will entail H as well. Probabilistic relations, however, can be

nonmonotonic: H might be highly probable given E, but improbable given $E \& E'$. For this reason, it's important for an agent assigning credences on the basis of her evidence to consider *all* of that evidence, and not simply draw conclusions from a subset of it. Carnap (1947) articulated the **Principle of Total Evidence** that a rational agent's credence distribution takes into account all of the evidence she possesses.

An agent may violate the Principle of Total Evidence by failing to take into account the *manner* in which she gained particular information. If the agent is aware of the mechanism by which a piece of information was received, it can be important to recognize facts about that mechanism as a component of her total evidence (along with the information itself). In Eddington's (1939) classic example, you draw a sample of fish from a lake, and all the fish you draw are longer than six inches. Normally, updating on this information would increase your confidence that every fish in the lake is at least that long. But if you know the net used to draw the sample has big holes through which shorter fish fall, a confidence increase is unwarranted. Here it's important to conditionalize not only on the lengths of the fish but also on how they were caught. The method by which your sample was collected has introduced an **observation selection effect** into the data.[10]

Observation selection effects are crucial to a famously counterintuitive probability puzzle, the **Monty Hall Problem** (Selvin 1975):

> In one of the games played on *Let's Make a Deal*, a prize is randomly hidden behind one of three doors. The contestant selects one door, then the host (Monty Hall) opens one of the doors the contestant didn't pick. Monty knows where the prize is, and makes sure to always open a door that doesn't have the prize behind it. (If both the unselected doors are empty, he randomly chooses which one to open.) After he opens an empty door, Monty asks the contestant if she wants what's behind the door she initially selected, or what's behind the other remaining closed door. Assuming she understands the details of Monty's procedure, how confident should the contestant be that the door she initially selected contains the prize?

Most people's initial reaction is to answer 1/2: the contestant originally spread her credence equally among the three doors; one of them has been revealed to be empty; so she should be equally confident that the prize is behind each of the remaining two. This analysis is illustrated by the following probability table:

	cr_1	cr_2
Prize behind door A	1/3	1/2
Prize behind door B	1/3	0
Prize behind door C	1/3	1/2

Here we've used the obvious partition of three locations where the prize might be. Without loss of generality, I've imagined that the contestant initially selects door A and Monty then opens door B. At time t_1—after the contestant has selected door A but before Monty has opened anything—she is equally confident that the prize is hidden behind each of the three doors. When Monty opens door B at t_2, the contestant should conditionalize on the prize's not being behind that door. This yields the cr_2 distribution listed above, which matches most people's intuitions.

Yet the contestant's *total* evidence at t_2 includes not only the fact that the prize isn't behind door B but also the fact that Monty chose that door to open for her. These two propositions aren't equivalent among the agent's doxastically possible worlds at t_1; there are possible worlds consistent with what the contestant knows about Monty's procedure in which door B is empty yet Monty opens door C. That door B was not only empty but was *revealed* to be empty is not expressible in the partition used above. So we need a richer partition, containing information both about the location of the prize and about what Monty does:

	cr_1	cr_2
Prize behind door A & Monty reveals B	1/6	1/3
Prize behind door A & Monty reveals C	1/6	0
Prize behind door B & Monty reveals C	1/3	0
Prize behind door C & Monty reveals B	1/3	2/3

Given what the agent knows of Monty's procedure, these four propositions partition her doxastic possibilities at t_1. At that time she doesn't know where the prize is, but she has initially selected door A (and Monty hasn't opened anything yet). If the prize is indeed behind door A, Monty randomly chooses whether to open B or C. So the contestant divides her 1/3 credence that the prize is behind door A equally between those two options. If the prize is behind door B, Monty is forbidden to open that door as well as the door the contestant selected, so Monty must open C. Similarly, if the prize is behind door C, Monty must open B.

At t_2 Monty has opened door B, so the contestant conditionalizes by setting her credences in the second and third partition members to 0, then multiplying the remaining values by a constant so that they sum to 1. This maintains the ratio between her credences on the first and fourth lines; initially she was twice as confident of the fourth as the first, so she remains twice as confident after the update. She is now 2/3 confident that the prize isn't behind the door she initially selected, and 1/3 confident that her initial selection was correct. If she wants the prize, the contestant should switch doors.

This is the correct analysis. If you find that surprising, the following explanation may help: When the contestant originally selected her door, she was 1/3 confident that the prize was behind it and 2/3 confident that the prize was somewhere else. If her initial pick was correct, she claims the prize just in case she sticks with that pick. But if her initial selection was wrong, she wins by switching to the other remaining closed door, because it must contain the prize. So there's a 1/3 chance that sticking is the winning strategy, and a 2/3 chance that switching will earn her the prize. Clearly switching is a better idea!

When I first heard the Monty Hall Problem, even that explanation didn't convince me. I only became convinced after I simulated the scenario over and over and found that sticking made me miss the prize roughly two out of three times. If you're not convinced, try writing a quick computer program or finding a friend with a free afternoon to act as Monty Hall for you a few hundred times. You'll eventually find that the probability table taking *total* evidence into account provides the correct analysis.[11]

One final note about total evidence: I may have convinced you that taking your total evidence into account is a good idea, but you might be concerned that it's impossible. After all, at each conscious moment an agent receives torrents of information from her environment. How can she take it *all* into account when assigning a credence to a particular proposition—say, the proposition that the cheese sandwich on the counter in front of her has not yet gone bad?

The nonmonotonicity of probabilistic relations means that a rational agent cannot afford to ignore any of her evidence. But many of the propositions an agent learns in a given moment will be *irrelevant* to the matter under consideration relative to her current credence distribution. That is, for many pieces of evidence her credence in the proposition at issue would be the same whether she conditionalized on that particular piece of evidence or not. As the agent ponders her cheese sandwich, information about the color of the bird that just flew by or the current position of her right hand makes no difference to her credence that the sandwich is edible. So while a rational agent doesn't *ignore* any of her total evidence, the irrelevance of much of that evidence permits her to focus in on the few pieces of evidence that are relevant to the proposition under consideration. For this reason, Bayesians often bypass discussion of an agent's total evidence in favor of discussing her total *relevant* evidence.[12]

4.3 Priors and standards

4.3.1 Initial priors

Consider a rational agent with probabilistic credences who updates by Conditionalization each time she gains new evidence, for her entire life. At a given moment t_i she has a credence distribution cr_i. She then gains new evidence E and updates by Conditionalization. Her unconditional cr_i values provide the priors for that update, and her cr_i values conditional on E provide the posteriors. By Conditionalization, these posteriors become her unconditional credences at the next time, t_j. Then she receives a new piece of evidence E'. Her unconditional cr_j values supply the priors for a new update, and her cr_j values conditional on E' are the posteriors.

And so it goes. We have already seen that if this agent updates by Conditionalization *every* time she learns something new, she will gain contingent certainties over time and never lose any of them. So her entire doxastic life will be a process of accumulating empirical evidence from her environment, building a snowball of information that never loses any of its parts.

What happens if we view that process backwards, working from the agent's present doxastic state back through the states she assigned in the past? Her current unconditional credences resulted from an earlier update by Conditionalization. Relative to that update, her current credences were the posteriors and some other distribution provided the priors. But those priors, in turn, came from a conditionalization. So they were once the posteriors of an even *earlier* set of priors. As we go backwards in time, we find a sequence of credence distributions, each of which was conditionalized to form the next. And since each conditionalization strictly added evidence, the earlier distributions contain successively less and less contingent information as we travel back.

Bayesian epistemologists often imagine marching backwards in this fashion until there's no farther back to go. They imagine that if you went back far enough, you would find a point at which the agent possessed literally no contingent information. This was the starting point from which she gained her

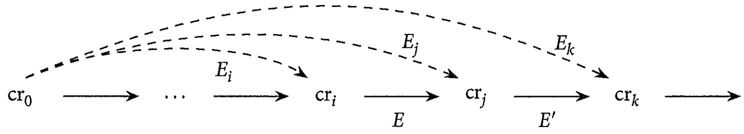

Figure 4.2 An initial prior?

very first piece of evidence, and made her first update by Conditionalization. The agent's credence distribution at this earliest point is sometimes called her **initial prior distribution** (or her "**ur-prior**").

Let's think about the properties an initial prior distribution would have. First, since the credence distributions that develop from an initial prior by Conditionalization are probability distributions, it's generally assumed that the initial prior satisfies the Kolmogorov axioms (and Ratio Formula) as well. Second, it's thought that since at the imagined initial moment (call it t_0) the agent possessed no contingent information, she should not have been certain of any contingent propositions. In other words, the initial prior distribution cr_0 should be regular (should assign nonextreme values to all contingent propositions). Finally, think about how cr_0 relates to a credence distribution our agent assigns at some arbitrary moment t_i later on. We could recover cr_i by conditionalizing cr_0 on the first piece of evidence the agent ever learned, then conditionalizing the result of that update on the second piece of evidence she learned, and so on until we reach cr_i. But since conditionalizing is cumulative, we could also roll up together all of these intermediate steps and get from cr_0 to cr_i in one move. Suppose the proposition E_i represents the agent's total evidence at t_i—the conjunction of all the individual pieces of evidence she's learned since t_0. Then as long as the agent has updated by conditionalizing at every step between t_0 and t_i, cumulativity guarantees that $\text{cr}_i(\cdot) = \text{cr}_0(\cdot \mid E_i)$. A rational agent's credence distribution at any given time is her initial prior distribution conditional on her total evidence at that time.

This idea is illustrated in Figure 4.2. Each credence distribution is generated from the previous one by conditionalizing on the evidence learned (solid arrows). But we can also derive each distribution directly (dashed arrows) by conditionalizing cr_0 on the agent's total evidence at the relevant time (E_i for cr_i, E_j for cr_j, etc.).

The initial-priors picture is an attractive one, and bears a certain mathematical elegance. The trouble is that it can at best be a myth. Was there ever a time in a real agent's life when she possessed *no* contingent information? Since cr_0 satisfies the probability axioms, it must be perfectly attuned to logical relations (such as mutual exclusivity and entailment) and assign a value of 1 to all tautologies. So the agent who assigns this initial prior must be omniscient logically while totally ignorant empirically. In seminars, David Lewis used to call such highly intelligent, blank creatures "superbabies"; while some Bayesian artificial intelligence systems may be like this, I doubt any human has ever been.[13] Moreover, I'm not sure it even makes sense for an agent with no contingent information to assign precise numerical credences to the kinds of elaborate, highly detailed empirical claims that are real humans' stock in trade.

4.3.2 Epistemic standards

Yet the formal mechanism employed by the initial priors myth—a regular probability distribution conditionalized on total evidence to generate credence distributions at arbitrary times—can be repurposed to represent something important in epistemology. To get a sense of what I'm talking about, let's consider an example:

Question: When playing a hand of five-card stud, how confident should you become that you'll win the hand upon learning that your last card will be the two of clubs?

Answer: Depends how far you are in the game, and what's happened up to that point.

Five-card stud is a poker game in which you receive a total of five cards, one at a time. Four of a kind (four out of five cards showing the same number) is an excellent, almost unbeatable hand in this game. So let's suppose that your first four cards in this particular hand of five-card stud were the jack of spades, the two of diamonds, the two of hearts, and then the two of spades. With that background information, discovering that your last card will be the two of clubs should make you almost certain that you'll win. (Depending in part on what you know of the other players' hands.)

An agent's **epistemic standards** govern how she reacts to particular pieces of news. These epistemic standards are determined in part by an agent's total evidence, and as such evolve over time. At the beginning of a hand of five-card stud, before any cards are dealt, learning that your last card will be the two of clubs (perhaps by peeking into the deck) would not make you very confident of winning the hand. Similarly, after seeing your first card (the jack of spades), a final two of clubs wouldn't seem like very good news. But for each successive two you receive after that point, your ongoing epistemic standards change such that learning the final card will be a two would make you more and more confident of a win.

When two people react differently to acquiring the same piece of information, they are applying different epistemic standards. We usually attribute the difference in their standards to differences in their previous experience. When one student in a class insists on answering every question, pontificates at length, and refuses to consider others' ideas, some of his fellow students might conclude that this person is the most knowledgeable in the room. But the teacher (or other students with more experience) might draw the opposite conclusion, informed by a broader pool of evidence about how particular personality types behave in conversation.

Yet how should we understand cases in which agents draw different conclusions despite sharing the same *total* evidence? Hiring committees form different beliefs about candidates' suitability from the same application files; jurors disagree about a defendant's guilt after witnessing the same trial; scientists embrace different hypotheses consistent with the same experimental data. These seem to be cases in which agents share a common body of total evidence, or at least total evidence *relevant* to the question at hand. So it can't be some further, unshared piece of extra evidence that's leading the agents to draw differing conclusions.

One could stubbornly maintain that in *every* real-life case in which agents interpret a piece of evidence differently, that difference is *entirely* attributable to the vagaries of their background information. But I think this would be a mistake. In addition to variations in their total evidence, agents have varying ways of interpreting their total evidence. Some people are naturally more skeptical than others, and so require more evidence to become confident of a particular proposition (that humans actually landed on the moon, that a lone gunman shot JFK, that the material world exists). Some people are interested in avoiding high confidence in falsehoods, while others are more interested in achieving high confidence in truths. (The former will tend to prefer noncommittal credences, while the latter will be more willing to adopt credence values near 0 and 1.) Some scientists are more inclined to believe elegant theories, while others incline toward the theory that hews closest to the data. (When the Copernican theory was first proposed, heliocentrism fit the available astronomical data *worse* than Ptolemaic approaches (Kuhn 1957).)

The five-card stud example illustrated what I will call *ongoing* epistemic standards. Ongoing epistemic standards reflect how an agent is disposed at a given time to assign attitudes in light of particular pieces of evidence she might receive. At any given time, an agent's credences can be determined from the last piece of evidence she acquired and the ongoing standards she possessed just before she acquired it (with the latter having been influenced by pieces of evidence she acquired even earlier than *that*). Yet there's another way to think about the influences on an agent's attitudes at a given time: we can separate out the influence of her *total* evidence from the influence of whatever additional, non-evidential factors dictate how she assesses that evidence. I refer to the latter as the agent's *ultimate* epistemic standards. An agent's ultimate epistemic standards capture her evidence-independent tendencies to respond to whatever package of total evidence might come her way.[14]

Warning

Don't take the talk of "standards" in "epistemic standards" too literally. It's not as if agents have a clearly defined set of *principles* that they apply to evidence whenever they interpret its significance, each of which is "an epistemic standard". Most of the time we draw conclusions from evidence without much deliberate thought at all, through a variety of methods that are scarcely systematic. Still, epistemologists need a way to refer to an agent's entire, disorganized bundle of dispositions to interpret evidence in particular ways—that bundle, taken as a unified whole, is what we mean by "her epistemic standards".[15]

In the next section we'll develop a formal, Bayesian representation of an agent's epistemic standards. Again, don't read the math as a report of what's going on in the agent's psychology—I don't do a bunch of probability calculations every time I try to figure out where my earbuds are. Instead, think of it as a succinct way to summarize what the agent would conclude from a variety of bodies of total evidence, regardless of how the agent actually goes about drawing such conclusions.

4.3.3 Hypothetical priors

How might we formally represent an agent's ultimate epistemic standards? A moment ago, I said that the attitudes an agent adopts at any given time combine two influences: her total evidence and her ultimate epistemic standards. So we can think of the agent's epistemic standards as a function from possible bodies of evidence to sets of attitudes adopted in response. This function can be represented in many ways—as a table listing outputs for specific inputs, as a complicated graph, etc. But Bayesians have a particularly useful representation already to hand. We can represent an agent's ultimate epistemic standards using a regular probability distribution Pr_H over her language $\mathcal{L}$, which we call her **hypothetical prior distribution**. I'll presently give an example of how to build such a representation from information about an agent's credences. The key point for now is that once we've constructed a hypothetical prior for an agent, we can recover from it her credence distribution cr_i at any given time t_i, as long as we know her total evidence E_i at that time. We simply conditionalize Pr_H on E_i, and the result is cr_i. Combining the agent's ultimate epistemic standards (Pr_H) with her total evidence (E_i) recovers her set of attitudes (cr_i).

Once we've constructed a hypothetical prior Pr_H for an agent, we can conditionalize it not only on bodies of total evidence she's possessed in the past but also on bodies of evidence she might possess in the future. The hypothetical prior is thus a highly efficient summary of all the attitudes an agent might ever adopt in response to bodies of evidence at various times of her life. Almost miraculously, such a compact summary will be available for any agent whose lifelong credences satisfy the five core Bayesian rules. This is guaranteed by the

Hypothetical Priors Theorem: Given any finite series of credence distributions $\{\text{cr}_1, \text{cr}_2, \ldots, \text{cr}_n\}$, each of which satisfies the probability axioms and Ratio Formula, let E_i be a conjunction of the agent's total evidence at t_i. If each cr_i is related to cr_{i+1} as specified by Conditionalization, then there exists at least one regular probability distribution Pr_H such that for all $1 \leqslant i \leqslant n$,

$$\text{cr}_i(\cdot) = \text{Pr}_H(\cdot \mid E_i)$$

In other words, if at each time in an agent's life her credence distribution satisfies the probability axioms and Ratio Formula, and if she updates those distributions from one time to the next according to Conditionalization, then there will necessarily exist at least one hypothetical prior Pr_H that relates to every credence distribution she ever assigns according to the formula above. (You'll prove this theorem in Exercise 4.8.)

Notice that the Hypothetical Priors Theorem specifies the existence of a *regular* probability distribution Pr_H. Given any rational Bayesian agent, we can recover her credence distribution at a given time by conditionalizing her hypothethical prior on her total contingent evidence at that time. Yet being regular, the hypothetical prior does not assign any contingent certainties itself.

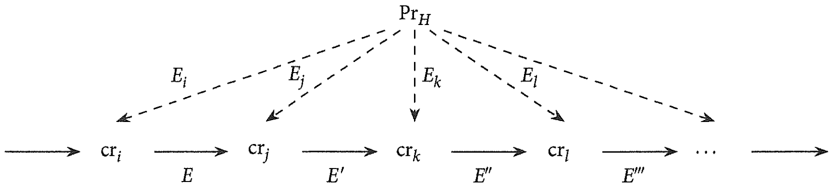

Figure 4.3 A hypothetical prior

So when we are confronted with the agent's credences at a particular time, we can cleanly factor out the two distinct influences on those credences: her total evidence (a set of contingent propositions in her language she takes for certain at that time) and her epistemic standards (represented by a hypothetical prior distribution assigning no contingent certainties). This is how a Bayesian represents the idea that ultimate epistemic standards are extra-evidential; instead of containing any contingent information about the world, epistemic standards express whatever the agent brings to bear on her evidence that isn't based on evidence itself.[16]

Hypothetical priors are convenient because they have a mathematical form with which we're already familiar: a regular probability distribution over language $\mathcal{L}$. Yet a hypothetical prior distribution is not a *credence* distribution.[17] An agent's hypothetical priors are not degrees of belief we imagine she espoused at some particular point in her life, or would espouse under some hypothetical conditions. This is what distinguishes them from the mythical initial priors.[18] Hypothetical priors summarize an agent's abstract evidential assessment tendencies, and stay constant throughout her life as long as she obeys the Conditionalization update rule. Instead of appearing somewhere *within* the series of credence distributions the agent assigns over time, the hypothetical prior "floats above" that series, combining with the agent's total evidence to create members of the series at each given time. As de Finetti puts it, "If we reason according to Bayes' theorem we do not change opinion. We keep the same opinion and we update it to the new situation" (de Finetti 1995, p. 100, translated by and quoted in Galavotti 2005, p. 215).

This arrangement is depicted in Figure 4.3. Again, the solid arrows represent conditionalizations from one time to the next, while the dashed arrows represent the possibility of generating an agent's distribution at any given time by conditionalizing Pr_H on her total evidence at that time.

Now let's show how we can construct a hypothetical prior from information about the credences an agent assigns over time. Suppose Ava has drawn two coins from a bin that contains only fair coins and coins biased toward heads.

Prior to time t_1 she has inspected both coins and determined them to be fair. Between t_1 and t_2 she flips the first coin, which comes up heads. Between t_2 and t_3, the second coin comes up tails.

Our language $\mathcal{L}$ will contain three atomic propositions: N, that neither coin Ava picked is biased; Ha, that the first flip comes up heads; and Hb, that the second flip is heads. Presumably the following probability table describes Ava's credences over time:

N	Ha	Hb	cr_1	cr_2	cr_3
T	T	T	1/4	1/2	0
T	T	F	1/4	1/2	1
T	F	T	1/4	0	0
T	F	F	1/4	0	0
F	T	T	0	0	0
F	T	F	0	0	0
F	F	T	0	0	0
F	F	F	0	0	0

In this example, Ava's total evidence at t_1 (or at least her total evidence representable in language $\mathcal{L}$) is N. We'll call this proposition E_1. Between t_1 and t_2, Ava learns Ha. So cr_2 is cr_1 conditionalized on Ha, and Ava's total evidence at t_2 (which we'll call E_2) is $N \,\&\, Ha$. After t_2 Ava learns $\sim Hb$, so cr_3 is cr_2 conditionalized on $\sim Hb$, and E_3 is $N \,\&\, Ha \,\&\, \sim Hb$. Notice that since N is part of Ava's evidence at all times reflected in this table, she assigns credence 0 throughout the table to any state-description on which N is false.

Since Ava's credence distributions cr_1 through cr_3 are probabilistic, and update by Conditionalization, the Hypothetical Priors Theorem guarantees the existence of at least one hypothetical prior Pr_H standing in a particular relation to Ava's credences. I've added a column to the probability table below representing one such hypothetical prior:

N	Ha	Hb	cr_1	cr_2	cr_3	Pr_H
T	T	T	1/4	1/2	0	1/16
T	T	F	1/4	1/2	1	1/16
T	F	T	1/4	0	0	1/16
T	F	F	1/4	0	0	1/16
F	T	T	0	0	0	21/64
F	T	F	0	0	0	11/64
F	F	T	0	0	0	11/64
F	F	F	0	0	0	5/64

As the Hypothetical Priors Theorem requires, Pr_H is a probability distribution (the values in the Pr_H column are non-negative and sum to 1), and it's regular (no contingent proposition receives a 0). Pr_H stands in the desired relationship to each of cr_1, cr_2, and cr_3: each of those distributions can be obtained from Pr_H by conditionalizing on Ava's total evidence at the relevant time. To take one example, consider cr_2. E_2 is $N \,\&\, Ha$. To conditionalize Pr_H on $N \,\&\, Ha$, we write a zero on each line whose state-description is inconsistent with $N \,\&\, Ha$. That puts zeroes on the third through eighth lines of the table. We then multiply the Pr_H values on the first and second lines by a constant (in this case, 8) so that the results sum to 1. This yields the cr_2 distribution in the table. With a bit of work you can verify that cr_1 results from conditionalizing Pr_H on E_1, and cr_3 is the result of conditionalizing Pr_H on E_3.

The hypothetical prior I wrote down isn't unique. I could have written down (infinitely) many other regular, probabilistic distributions that stand in the required relation to cr_1 through cr_3. This reveals that the information in the original table underdescribes Ava's ultimate epistemic standards, even over our fairly limited language $\mathcal{L}$. For instance, the original table doesn't tell us what credences Ava would've assigned had she learned before t_1 that at least one of the coins was biased. The Pr_H I've provided makes very specific assumptions about Ava's tendencies for that case. (For a fun exercise, figure out what that Pr_H assumes about the biased coins in the bin.) But I could've made different assumptions, and generated a different hypothetical prior consistent with cr_1 through cr_3.

Interestingly, those assumptions don't matter much for practical purposes. Suppose you're working with the series of credence distributions an agent assigned up to some designated time. Typically there will be multiple hypothetical priors consistent with that series of distributions. Those hypothetical priors will differ on what attitudes the agent would have assigned had she received different bodies of total evidence prior to the designated time. But given the course of evidence the agent actually received and series of credence distributions she actually assigned, every hypothetical prior consistent with that series of distributions will make the same predictions about how she will respond to particular pieces of evidence after the designated time. (Assuming she continues to conditionalize.) For instance, every hypothetical prior consistent with Ava's cr_1 and cr_2 distributions will predict the same response to her learning $\sim Hb$ between t_2 and t_3. So which of the available hypothetical priors we use to represent a particular agent's epistemic standards turns out to be irrelevant going forward.

On the other hand, when distinct *agents* have different hypothetical priors, those differences can be important. Plugging an agent's total evidence at a given time into her hypothetical prior yields her credence distribution at that time. When two agents have different hypothetical priors, plugging in the same body of total evidence may yield different results. So two agents may assign different credences to the same proposition in light of the same total evidence. The difference in hypothetical priors is a Bayesian's way of representing that these agents interpret evidence differently, and so may draw different conclusions from the same total body of information.

The obvious next question is whether they can both be *rational* in doing so. Evidence and epistemic standards come up in a variety of contexts in epistemology. As we've just seen, Bayesian epistemology provides an elegant formal apparatus for separating out these two elements. But once we've isolated them, the next question to ask about ultimate epistemic standards is whether anything goes. Is *any* hypothetical prior rational, so long as it's probabilistic? Some probabilistic hypothetical priors will be anti-inductive, or will recommend highly skeptical attitudes in the face of everyday bodies of total evidence. Can we rule out such hypothetical priors as rationally impermissible? Can we go even farther than that, laying down enough rational constraints on ultimate epistemic standards so that any time two agents interpret the same evidence differently, at least one of them must be interpreting it irrationally? This will be our first topic in Chapter 5, as we distinguish Objective from Subjective Bayesianism.

4.4 Exercises

Unless otherwise noted, you should assume when completing these exercises that the credence distributions under discussion satisfy the probability axioms and Ratio Formula. You may also assume that whenever a conditional credence expression occurs or a proposition is conditionalized upon, the needed proposition has nonzero unconditional credence so that conditional credences are well-defined.

Problem 4.1. Galileo intends to determine whether gravitational acceleration is affected by mass by dropping two cannonballs with differing masses off the Leaning Tower of Pisa. Conditional on there being no effect, he is 95% confident that the cannonballs will land within 0.1 seconds of each other. (The experiment isn't perfect—one ball might hit a bird.) Conditional on mass affecting acceleration, he is 80% confident that the balls *won't* land within 0.1 seconds of each other. (There's some chance that although mass affects acceleration, it doesn't have *much* of an effect.)[19]

(a) Before performing the experiment, Galileo is 30% confident that mass does not affect acceleration. How confident is he that the cannonballs will land within 0.1 seconds of each other?

(b) After Galileo conditionalizes on the evidence that the cannonballs landed within 0.1 seconds of each other, how confident is he in each hypothesis?

Problem 4.2. Prove that conditionalizing retains certainties. In other words, prove that if $\text{cr}_i(H) = 1$ and cr_j is generated from cr_i by Conditionalization, then $\text{cr}_j(H) = 1$ as well.

Problem 4.3. Prove that conditionalization is cumulative. That is, prove that for any cr_i, cr_j, and cr_k, conditions 1 and 2 below entail condition 3.

1. For any proposition X in $\mathcal{L}$, $\text{cr}_j(X) = \text{cr}_i(X \mid E)$.
2. For any proposition Y in $\mathcal{L}$, $\text{cr}_k(Y) = \text{cr}_j(Y \mid E')$.
3. For any proposition Z in $\mathcal{L}$, $\text{cr}_k(Z) = \text{cr}_i(Z \mid E \,\&\, E')$.

Problem 4.4.

(a) Provide a real-life example in which an agent conditionalizes on new evidence, yet her credence in a proposition *compatible with* the evidence decreases. That is, provide an example in which H and E are consistent, yet $\text{cr}_2(H) < \text{cr}_1(H)$ when E is learned between t_1 and t_2.

(b) Prove that when an agent conditionalizes on new evidence, her credence in a proposition that *entails* the evidence cannot decrease. That is, when $H \vDash E$, it must be the case that $\text{cr}_2(H) \geq \text{cr}_1(H)$ when E is learned between t_1 and t_2.

(c) Prove that as long as $\text{cr}_1(H)$ and $\text{cr}_1(E)$ are both nonextreme, conditionalizing on E increases the agent's credence in H when $H \vDash E$.[20]

Problem 4.5. Reread the details of the Base Rate Fallacy example in Section 4.1.2. After you apply the diagnostic test once and it yields a positive result, your odds for D should be $9 : 999$.

(a) Suppose you apply the test a second time to the same subject, and it yields a positive result once more. What should your odds for the subject's having the disease be now? (Assume that D screens off the results of the first test from the results of the second.)

(b) How many consecutive tests (each independent of all the prior test results conditional on both D and $\sim D$) would have to yield positive results before your confidence that the subject has the disease exceeded 50%?

(c) Does this shed any light on why patients diagnosed with rare diseases are often advised to seek a second opinion? Explain.

Problem 4.6. Your friend Jones is a gambler. He even gambles about whether to gamble! Every time he goes to the track, he flips a fair coin to determine whether to bet that day. If it comes up heads he bets on his favorite horse, Speedy. If it comes up tails he doesn't bet at all.

On your way to the track today, you were 1/6 confident that out of the six horses running, Speedy would win. You were 1/2 confident that Jones's coin would come up heads. And you considered the outcome of the horse race independent of the outcome of the coin flip. But then you saw Jones leaving the track with a smile on his face. The smile tells you that either Jones bet on Speedy and won, or Jones didn't bet and Speedy didn't win.[21]

(a) Using a language with the atomic propositions H (for heads on the coin) and S (for a Speedy win), express the information you learn when you see Jones smiling.

(b) After updating on this information by conditionalizing, how confident are you that Speedy won? How confident are you that the coin came up heads?

(c) Explain why one of the unconditional credences you calculated in part (b) differs from its prior value and the other one doesn't. Be sure to include an explanation of why *that* unconditional credence was the one that changed out of the two. ("Because that's what the math says" is not an adequate explanation—we want to know why the mathematical outcome *makes sense*.)

(d) Evaluate the following claim: "If an agent starts off viewing proposition X as irrelevant to proposition Y, then learning that X and Y have the same truth-value shouldn't change that agent's credence in Y." Explain your answer.

Problem 4.7. At t_1, t_2, and t_3, Jane assigns credences over the language $\mathcal{L}$ constructed from atomic propositions P and Q. Jane's distributions satisfy constraints 1 through 6:

1. At t_1, Jane is certain of $Q \supset P$, anything that proposition entails, and nothing else.

2. Between t_1 and t_2 Jane learns P and nothing else. She updates by conditionalizing between those two times.
3. $\text{cr}_1(Q \mid P) = 2/3$.
4. $\text{cr}_3(Q \mid {\sim}P) = 1/2$.
5. $\text{cr}_3(P \supset Q) = \text{cr}_2(P \supset Q)$.
6. At t_3, Jane is certain of ${\sim}(P \,\&\, Q)$, anything that proposition entails, and nothing else.

(a) Completely specify Jane's credence distributions at t_2 and t_3.

(b) Create a hypothetical prior for Jane. In other words, specify a regular probabilistic distribution Pr_H over $\mathcal{L}$ such that cr_1 can be generated from Pr_H by conditionalizing on Jane's set of certainties at t_1; cr_2 is Pr_H conditionalized on Jane's certainties at t_2; and cr_3 is Pr_H conditionalized on Jane's t_3 certainties.

(c) Does Jane update by Conditionalization between t_2 and t_3? Explain how you know.

(d) The Hypothetical Priors Theorem says that *if* an agent always updates by conditionalizing, *then* her credences can be represented by a hypothetical prior distribution. Is the converse of this theorem true?

Problem 4.8. In this exercise we'll prove the Hypothetical Priors Theorem, by showing how to construct the needed Pr_H for any given series of credences.

Start by supposing we have a finite series of credence distributions $\{\text{cr}_1, \text{cr}_2, \ldots, \text{cr}_n\}$, each of which satisfies the probability axioms and Ratio Formula. Suppose also that each cr_i is related to cr_{i+1} as specified by Conditionalization. Then, for each cr_i, define E_i as a proposition logically equivalent to the conjunction of all the propositions receiving cr_i values of 1.[22]

(a) Let's start by focusing only on cr_1, the first credence distribution in the series. Describe how, given cr_1, to construct a regular probability distribution Pr_H such that cr_1 results from conditionalizing Pr_H on E_1. (It might help here to review the Ava example from page 111ff.)

(b) Now, using the cumulativity of conditionalization, explain why the fact that cr_1 results from conditionalizing Pr_H on E_1 guarantees that for every cr_i in the series, $\text{cr}_i(\cdot) = \text{Pr}_H(\cdot \mid E_i)$.

Problem 4.9. Suppose you have a finite partition $\{B_1, B_2, \ldots B_n\}$ of propositions from $\mathcal{L}$. Suppose also that between t_1 and t_2 you conditionalize on evidence equivalent to a disjunction of some of the Bs. Prove that for any A in $\mathcal{L}$ and any B_i such that $\text{cr}_2(B_i) > 0$,

$$\mathrm{cr}_2(A \mid B_i) = \mathrm{cr}_1(A \mid B_i)$$

Problem 4.10. Do you think only one set of ultimate epistemic standards is rationally permissible? Put another way: If two agents' series of credence distributions cannot be represented by the same hypothetical prior distribution, must at least one of them have assigned irrational credences at some point?

4.5 Further reading

INTRODUCTIONS AND OVERVIEWS

Ian Hacking (2001). *An Introduction to Probability and Inductive Logic.* Cambridge: Cambridge University Press

Chapter 15 works through many excellent examples of applying Bayes's Theorem to manage complex updates.

CLASSIC TEXTS

Rudolf Carnap (1947). On the Application of Inductive Logic. *Philosophy and Phenomenological Research* 8, pp. 133–48

Section 3 contains Carnap's original discussion of the Principle of Total Evidence.

EXTENDED DISCUSSION

Paul Teller (1973). Conditionalization and Observation. *Synthese* 26, pp. 218–58

Offers a number of arguments for the Conditionalization updating norm. (In Chapter 9 we'll discuss the Dutch Book argument for Conditionalization that Teller provides.)

Isaac Levi (1980). *The Enterprise of Knowledge.* Boston: The MIT Press

Though Levi's notation and terminology are somewhat different from mine, Chapter 4 thoroughly works through the mathematics of hypothetical priors.

Levi also discusses various historically important Bayesians' positions on how many distinct hypothetical priors are rationally permissible.

Christopher J.G. Meacham (2016). Ur-Priors, Conditionalization, and Ur-Prior Conditionalization. *Ergo* 3, pp. 444–92

Meacham considers a number of possible interpretations of hypothetical priors, and how they might be linked to an agent's credences at specific times by Conditionalization.

Notes

1. Remember that 1 is not a prime number, while 2 is!
2. A bit of reflection on Equation (4.2) will reveal that the constant by which we multiply in the second step of our process—the **normalization factor**—is always the reciprocal of the agent's initial unconditional credence in the evidence. In other words, the second step *divides* all nonzero state-description credences by $\text{cr}_i(E)$.
3. We can also now see an alternate explanation for steps (3.51) and (3.53) of Lewis's triviality proof from Section 3.3. The proposal assessed there is that for some conditional $\rightarrow$, the agent's conditional credence $\text{cr}(Z \mid Y)$ for any Y and Z in $\mathcal{L}$ equals her unconditional credence in $Y \rightarrow Z$. Whatever motivates that proposal, we should want the proposal to remain true even after the agent learns some information X. If the relevant values are going to match after conditionalization on X, it must be true before conditionalization that $\text{cr}(Y \rightarrow Z \mid X) = \text{cr}(Z \mid Y \,\&\, X)$, which is just Equation (3.59).
4. Thanks to Joel Velasco for discussion, and for the example.
5. For reasons we are now in a position to understand, the term "posterior" is sometimes used ambiguously in the Bayesian literature. I have defined "posterior" as an agent's conditional credence in the hypothesis given the evidence—$\text{cr}(H \mid E)$. If the agent updates by conditionalizing on E, this will equal her unconditional credence in the hypothesis after the update. The terms "prior" and "posterior" come from the fact that on an orthodox Bayesian position, those quantities pick out the agent's unconditional credences in the hypothesis before and after the update. But unorthodox Bayesians who prefer an alternative updating rule to Conditionalization nevertheless sometimes refer to an agent's post-update credence in a hypothesis as her "posterior". As I've defined the term, this is strictly speaking incorrect.
6. An acquaintance involved with neuroscientific research once claimed that when a prisoner in the American penal system comes up for parole, a particular kind of brain scan can predict with greater than 90% accuracy whether that prisoner will, if released, be sent back to jail within a specified period of time. He suggested that we use this brain scan in place of traditional parole board hearings, whose predictive accuracy is much lower. To get at some of my ethical concerns with this proposal, I asked why we don't just apply the brain scan to everyone in America, rather than waiting to see if a

person winds up in jail. He replied that the base rates make this impractical: While the recidivism rate among prisoners is fairly high, the percentage of Americans who wind up in jail is much lower, so the scan would generate far too many false positives if used on the general population.

7. In Section 2.2.3 I mentioned that Bayesians often work with an agent's set of doxastically possible worlds instead of the full set of logically possible worlds, understanding "mutually exclusive" and "tautology" in the Kolmogorov axioms in terms of this restricted doxastic set. The Regularity Principle concerns the *full* set of logically possible worlds—it forbids assigning credence 0 to any proposition that is true in at least one of them. So for the rest of this section, references to "contingent propositions", "tautologies", etc. should be read against that full logical set of possibilities.
8. What most of us call the "Regularity Principle" Dennis Lindley dubbed "Cromwell's Rule". He wrote, "A simple result that follows from Bayes' theorem is that it is inadvisable to attach probabilities of zero to uncertain events, for if the prior probability is zero so is the posterior, whatever be the data.... In other words, if a decision-maker thinks something cannot be true and interprets this to mean it has zero probability, he will never be influenced by *any* data, which is surely absurd. So leave a little probability for the moon being made of green cheese; it can be as small as 1 in a million, but have it there since otherwise an army of astronauts returning with samples of the said cheese will leave you unmoved.... As Oliver Cromwell told the Church of Scotland, 'I beseech you, in the bowels of Christ, think it possible you may be mistaken'" (Lindley 1985, p. 104, emphasis in original). (Thanks to Patrick Cronin for bringing this to my attention and to Wikipedia for the reference.)
9. Throughout this section I identify credence 1 with absolute certainty in a proposition and credence 0 with ruling that proposition out. This becomes more complicated when we consider events with infinitely many possible outcomes; the relevant complications will be addressed in Chapter 5.
10. Observation selection effects pop up all over the place in real life—perhaps you think the refrigerator light is *always* on because it's on whenever you open the door to look. Here's my favorite example: During World War II, the American military showed mathematician Abraham Wald data indicating that planes returning from dogfights had more bullet holes in the fuselage than in the engine. The military was considering shifting armor from the engine to the fuselage. Wald recommended exactly the opposite, on grounds that it was the *returning* planes that had holes in the fuselage but not the engines (Wainer 2011, recounted in Ellenberg 2014, pp. 12–3).
11. A similar situation occurs in Bradley (2010): Colin Howson argued that a so-called "Thomason case" provides a counterexample to Conditionalization. Bradley replies that if we analyze the agent's *total* evidence in the case—including evidence about how he came to have his evidence—the supposed counterexample disappears.
12. You may have noticed that in the Monty Hall Problem, accounting for the agent's total relevant evidence required us to move from a coarser-grained partition of her doxastic possibilities (Prize behind door A/B/C) to a finer-grained partition (Prize behind A & Monty reveals B, Prize behind A & Monty reveals C, etc.). Whether a conditionalization yields the right results often depends on the richness of the language in which the agent represents her doxastic possibilities; a language without enough detail may miss aspects

of her total relevant evidence. For more on selecting an appropriately detailed language for updating, and some formal results on how one can know when one's language is rich enough, see Titelbaum (2013a, Ch. 8).

13. I learned of Lewis's "superbaby" talk from Alan Hájek. Susan Vineberg suggested to me that Lewis's inspiration for the term may have been I.J. Good's (1968) discussion of "an infinitely intelligent newborn baby having built-in neural circuits enabling him to deal with formal logic, English syntax, and subjective probability"—a discussion to which we shall return in Chapter 6.
14. The earliest sustained usage I can find of the phrase "epistemic standards" occurs in Foley (1987, p. 33ff.). While Foley may have coined the *name*, he wasn't the first to discuss the *notion* I'm calling "ultimate epistemic standards"—Levi (1980) has a similar notion of "confirmational commitments", and in Chapter 6 we'll see Carnap proposing a similar thing much earlier. Levi's discussion is particularly important because it lays out the mathematical formalism for epistemic standards I'm about to present.

    While we're discussing terminology: Given the choice, I'd prefer the term "evidential standards" to "epistemic standards", because the former emphasizes the standards' taking evidence as their inputs, while the latter introduces an unwanted association with knowledge. Yet the "epistemic standards" terminology is so entrenched at this point that I've given up fighting against it. (Thanks to Laura Callahan for the pointer to Foley.)
15. Compare: When we speak admiringly of someone's "high moral standards", I don't think we're necessarily attributing the possession of systematic moral principles. This way of talking seems as compatible with virtue ethics as any other moral approach.
16. Catrin Campbell-Moore pointed out to me that the Hypothetical Priors Theorem assumes that a contingent proposition receives credence 1 at t_i only if it is entailed by E_i. If this assumption fails, we will still be able find at least one hypothetical prior Pr_H for the series of credence distributions, but no *regular* hypothetical prior will be available.

    The assumption reflects the idea that the only thing that can justify certainty in a contingent proposition is empirical evidence. It's interesting to think about cases in which this assumption might fail. It might be that particular propositions, despite being logically contingent, nevertheless merit rational a priori certainty. Or another way the assumption could fail is if credence 1 doesn't always represent certainty. We'll see examples where this might occur in Section 5.4.
17. This is why defining hypothetical priors as regular does not commit us to the Regularity Principle as a constraint on rational agents' credences. We make hypothetical priors regular so they will remain independent of contingent evidence, and therefore capable of representing extra-evidential influences on an agent's attitudes. But since hypothetical priors aren't credences, this doesn't say anything about whether credences should be regular or not. Moreover, the Hypothetical Priors Theorem applies only to agents who update by Conditionalization, while Conditionalization conflicts with the Regularity Principle.
18. In the Bayesian literature, the terms "initial prior", "ur-prior", and "hypothetical prior" are often used interchangeably. To me, the former two connote that the prior was assigned by the agent at some early time. So I've selected "hypothetical prior" to emphasize the use of a mathematical representation that does not correspond to any credences the agent ever actually assigned. Unfortunately, the term "hypothetical prior"

has also been used for a very specific notion within the literature on the Problem of Old Evidence (as in Bartha and Hitchcock 1999, p. S349). Here I simply note the distinction between that usage and the one I intend; I'll explain the alternate usage in Section 12.1.3.

19. This is a version of a problem from Julia Staffel.
20. This problem was inspired by a problem of Sarah Moss's.
21. This story is adapted from Hart and Titelbaum (2015).
22. We define E_i this way for reasons discussed in note 16.

Bibliography of Volumes 1 & 2

Achinstein, Peter (1963). Variety and Analogy in Confirmation Theory. *Philosophy of Science* 3, pp. 207–21.

Adams, Ernest (1962). On Rational Betting Systems. *Archiv für mathematische Logik und Grundlagenforschung* 6, pp. 7–29.

Adams, Ernest (1965). The Logic of Conditionals. *Inquiry* 8, pp. 166–97.

Alchourrón, Carlos E., Peter Gärdenfors, and David Makinson (1985). On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *The Journal of Symbolic Logic* 50, pp. 510–30.

Allais, Maurice (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulates et axiomes de l'ecole Américaine. *Econometrica* 21, pp. 503–46.

Amrhein, Valentin, Sander Greenland, and Blake McShane (2019). Scientists Rise up against Statistical Significance. *Nature* 567, pp. 305–7.

Armendt, Brad (1980). Is There a Dutch Book Argument for Probability Kinematics? *Philosophy of Science* 47, pp. 583–8.

Armendt, Brad (1992). Dutch Strategies for Diachronic Rules: When Believers See the Sure Loss Coming. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1, pp. 217–29.

Arntzenius, Frank (1993). The Common Cause Principle. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 2, pp. 227–37.

Arntzenius, Frank (2003). Some Problems for Conditionalization and Reflection. *The Journal of Philosophy* 100, pp. 356–70.

Arrow, Kenneth J. (1951). *Social Choice and Individual Values.* New York: John Wiley and Sons.

Bandyopadhyay, Prasanta S. and Malcolm R. Forster (2011). *Philosophy of Statistics.* Vol. 7. Handbook of the Philosophy of Science. Amsterdam: Elsevier.

Barnett, Jeffrey A. (1981). Computational Methods for a Mathematical Theory of Evidence. In: *Proceedings of the 7th International Joint Conference on AI,* pp. 868–75.

Bartha, Paul and Christopher R. Hitchcock (1999). No One Knows the Date or the Hour: An Unorthodox Application of Rev. Bayes's Theorem. *Philosophy of Science* 66, S339–53.

Berger, James O. and Donald A. Berry (1988). Statistical Analysis and the Illusion of Objectivity. *American Scientist* 76, pp. 159–65.

Bergmann, Merrie, James Moor, and Jack Nelson (2013). *The Logic Book.* 6th edition. New York: McGraw Hill.

Berker, Selim (2013). Epistemic Teleology and the Separateness of Propositions. *Philosophical Review* 122, pp. 337–93.

Bernoulli, Daniel (1738/1954). Exposition of a New Theory on the Measurement of Risk. *Econometrica* 22, pp. 23–36.

Bernoulli, Jacob (1713). *Ars Conjectandi.* Basiliae.

Bertrand, Joseph (1888/1972). *Calcul des probabilités.* 2nd edition. New York: Chelsea Publishing Company.

Bickel, P.J., E.A. Hammel, and J.W. O'Connell (1975). Sex Bias in Graduate Admissions: Data from Berkeley. *Science* 187, pp. 398–404.

Birnbaum, Allan (1962). On the Foundations of Statistical Inference. *Journal of the American Statistical Association* 57, pp. 269–306.

Bolzano, Bernard (1837/1973). *Wissenschaftslehre*. Translated by Jan Berg under the title *Theory of Science*. Dordrecht: Reidel.

Bovens, Luc and Stephan Hartmann (2003). *Bayesian Epistemology*. Oxford: Oxford University Press.

Bradley, Darren (2010). Conditionalization and Belief *De Se*. *Dialectica* 64, pp. 247–50.

Bradley, Darren (2011). Self-location Is No Problem for Conditionalization. *Synthese* 182, pp. 393–411.

Bradley, Darren (2015). *A Criticial Introduction to Formal Epistemology*. London: Bloomsbury.

Bratman, Michael E. (1987). *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.

Brier, George (1950). Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review* 78, pp. 1–3.

Briggs, R.A. (2010). Putting a Value on Beauty. In: *Oxford Studies in Epistemology*. Ed. by Tamar Szabó Gendler and John Hawthorne. Vol. 3. Oxford University Press, pp. 3–34.

Briggs, R.A. and Richard Pettigrew (2020). An Accuracy-Dominance Argument for Conditionalization. *Noûs* 54, pp. 162–81.

Broome, John (1999). Normative Requirements. *Ratio* 12, pp. 398–419.

Buchak, Lara (2013). *Risk and Rationality*. Oxford: Oxford University Press.

Campbell-Moore, Catrin and Benjamin A. Levinstein (2021). Strict Propriety Is Weak. *Analysis* 81.1, pp. 8–13.

Capotorti, Andrea and Barbara Vantaggi (2000). Axiomatic Characterization of Partial Ordinal Relations. *International Journal of Approximate Reasoning* 24, pp. 207–19.

Carnap, Rudolf (1945). On Inductive Logic. *Philosophy of Science* 12, pp. 72–97.

Carnap, Rudolf (1947). On the Application of Inductive Logic. *Philosophy and Phenomenological Research* 8, pp. 133–48.

Carnap, Rudolf (1950). *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Carnap, Rudolf (1955/1989). Statistical and Inductive Probability. In: *Readings in the Philosophy of Science*. Ed. by Baruch A. Brody and Richard E. Grandy. 2nd edition. Hoboken: Prentice-Hall, pp. 279–87.

Carnap, Rudolf (1962a). *Logical Foundations of Probability*. 2nd edition. Chicago: University of Chicago Press.

Carnap, Rudolf (1962b). The Aim of Inductive Logic. In: *Logic, Methodology, and the Philosophy of Science*. Ed. by P. Suppes, E. Nagel, and A. Tarski. Stanford University: Stanford University Press, pp. 303–18.

Carnap, Rudolf (1971). A Basic System of Inductive Logic, Part 1. In: *Studies in Inductive Logic and Probability*. Ed. by Rudolf Carnap and Richard C. Jeffrey. Vol. I. Berkeley: University of California Press, pp. 33–166.

Carnap, Rudolf (1980). A Basic System of Inductive Logic, Part 2. In: *Studies in Inductive Logic and Probability*. Ed. by Richard C. Jeffrey. Vol. II. Berkeley: University of California Press, pp. 7–156.

Carr, Jennifer (2017). Epistemic Utility Theory and the Aim of Belief. *Philosophy and Phenomenological Research* 95, pp. 511–34.

Cartwright, Nancy (1979). Causal Laws and Effective Strategies. *Noûs* 13, pp. 419–37.
Carver, Ronald P. (1978). The Case against Statistical Significance Testing. *Harvard Educational Review* 48, pp. 378–99.
Chandler, Jake (2013). Contrastive Confirmation: Some Competing Accounts. *Synthese* 190, pp. 129–38.
Cherniak, Christopher (1986). *Minimal Rationality.* Cambridge, MA: The MIT Press.
Chihara, C. (1981). Quine and the Confirmational Paradoxes. In: *Midwest Studies in Philosophy 6: Foundations of Analytic Philosophy.* Ed. by P. French, H. Wettstein, and T. Uehling. Minneapolis: University of Minnesota Press, pp. 425–52.
Christensen, David (1991). Clever Bookies and Coherent Beliefs. *The Philosophical Review* 100, pp. 229–47.
Christensen, David (1999). Measuring Confirmation. *The Journal of Philosophy* 96, pp. 437–61.
Christensen, David (2001). Preference-Based Arguments for Probabilism. *Philosophy of Science* 68, pp. 356–76.
Christensen, David (2004). *Putting Logic in its Place.* Oxford: Oxford University Press.
Colyvan, Mark (2004). The Philosophical Significance of Cox's Theorem. *International Journal of Approximate Reasoning* 37, pp. 71–85.
Cox, Richard T. (1946). Probability, Frequency and Reasonable Expectation. *American Journal of Physics* 14, pp. 1–13.
Cox, Richard T. (1961). *The Algebra of Probable Inference.* Baltimore, MD: The Johns Hopkins Press.
Crupi, Vincenzo, Branden Fitelson, and Katya Tentori (2008). Probability, Confirmation, and the Conjunction Fallacy. *Thinking & Reasoning* 14, pp. 182–99.
Crupi, Vincenzo, Katya Tentori, and Michel Gonzalez (2007). On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues. *Philosophy of Science* 74, pp. 229–52.
Davidson, Donald (1984). *Inquiries into Truth and Interpretation.* Oxford: Clarendon Press.
Davidson, Donald, J.C.C. McKinsey, and Patrick Suppes (1955). Outlines of a Formal Theory of Value, I. *Philosophy of Science* 22, pp. 14–60.
de Finetti, Bruno (1931/1989). Probabilism: A Critical Essay on the Theory of Probability and the Value of Science. *Erkenntnis* 31, pp. 169–223. Translation of B. de Finetti, *Probabilismo,* Logos 14, pp. 163–219.
de Finetti, Bruno (1937/1964). Foresight: Its Logical Laws, its Subjective Sources. In: *Studies in Subjective Probability.* Ed. by Henry E. Kyburg Jr and H.E. Smokler. New York: Wiley, pp. 94–158. Originally published as "La prévision; ses lois logiques, ses sources subjectives" in *Annales de l'Institut Henri Poincaré* 7, pp. 1–68.
de Finetti, Bruno (1949/1951). La 'Logica del Plausible' Secondo la Conçezione di Polya. *Atti della XLII Riunione, Societa Italiana per il Progresso delle Scienze.* Presented in 1949, published in 1951, pp. 227–36.
de Finetti, Bruno (1974). *Theory of Probability.* Vol. 1. New York: Wiley.
de Finetti, Bruno (1995). *Filosofia della probabilità.* Ed. by Alberto Mura. Milan: Il Saggiatore.
DeGroot, Morris H. (1970). *Optimal Statistical Decisions.* Hoboken, New Jersey: Wiley.
Dempster, Arthur P. (1966). New Methods for Reasoning towards Posterior Distributions Based on Sample Data. *Annals of Mathematical Statistics* 37, pp. 355–74.
Dickson, Michael and Davis Baird (2011). Significance Testing. In: *Philosophy of Statistics.* Ed. by Prasanta S. Bandyopadhyay and Malcolm R. Forster. Vol. 7. Handbook of the Philosophy of Science. Amsterdam: Elsevier.

Domotor, Zoltan (1969). *Probabilistic Relational Structures and their Applications.* Technical Report 144. Stanford University: Institute for Mathematical Studies in the Social Sciences.

Earman, John (1992). *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory.* Cambridge, MA: The MIT Press.

Easwaran, Kenny (2013). Expected Accuracy Supports Conditionalization—and Conglomerability and Reflection. *Philosophy of Science* 80, pp. 119–42.

Easwaran, Kenny (2014a). Decision Theory without Representation Theorems. *Philosophers' Imprint* 14, pp. 1–30.

Easwaran, Kenny (2014b). Regularity and Hyperreal Credences. *Philosophical Review* 123, pp. 1–41.

Eddington, A. (1939). *The Philosophy of Physical Science.* Cambridge: Cambridge University Press.

Edwards, A.W.F. (1972). *Likelihood: An Account of the Statistical Concept of Likelihood and its Application to Scientific Inference.* Cambridge: Cambridge University Press.

Eells, Ellery (1982). *Rational Decision and Causality.* Cambridge Studies in Philosophy. Cambridge: Cambridge University Press.

Eells, Ellery (1985). Problems of Old Evidence. *Pacific Philosophical Quarterly* 66, pp. 283–302.

Eells, Ellery and Branden Fitelson (2000). Measuring Confirmation and Evidence. *Journal of Philosophy* 97, pp. 663–72.

Eells, Ellery and Branden Fitelson (2002). Symmetries and Asymmetries in Evidential Support. *Philosophical Studies* 107, pp. 129–42.

Efron, B. (1986). Why Isn't Everyone a Bayesian? *The American Statistician* 40, pp. 1–5.

Egan, Andy (2007). Some Counterexamples to Causal Decision Theory. *Philosophical Review* 116, pp. 93–114.

Egan, Andy and Michael G. Titelbaum (2022). Self-locating Belief. Forthcoming in *The Stanford Encyclopedia of Philosophy.*

Elga, Adam (2000). Self-locating Belief and the Sleeping Beauty Problem. *Analysis* 60, pp. 143–7.

Elga, Adam (2007). Reflection and Disagreement. *Noûs* 41, pp. 478–502.

Elga, Adam (2010). Subjective Probabilities Should Be Sharp. *Philosophers' Imprint* 10, pp. 1–11.

Ellenberg, Jordan (2014). *How Not to Be Wrong: The Power of Mathematical Thinking.* New York: Penguin Press.

Ellis, Robert Leslie (1849). On the Foundations of the Theory of Probabilities. *Transactions of the Cambridge Philosophical Society* VIII, pp. 1–6.

Ellsberg, Daniel (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics* 75, pp. 643–69.

Feldman, Richard (2007). Reasonable Religious Disagreements. In: *Philosophers without Gods: Meditations on Atheism and the Secular Life.* Ed. by Louise M. Antony. Oxford: Oxford University Press, pp. 194–214.

Feller, William (1968). *An Introduction to Probability Theory and its Applications.* 3rd edition. New York: Wiley.

Fermat, Pierre and Blaise Pascal (1654/1929). Fermat and Pascal on Probability. In: *A Source Book in Mathematics.* Ed. by D. Smith. Translated by Vera Sanford. New York: McGraw-Hill, pp. 546–65.

Fine, Terrence L. (1973). *Theories of Probability: An Examination of Foundations.* New York, London: Academic Press.

Fishburn, Peter C. (1981). Subjective Expected Utility: A Review of Normative Theories. *Theory and Decision* 13, pp. 129–99.

Fishburn, Peter C. (1986). The Axioms of Subjective Probability. *Statistical Science* 1, pp. 335–45.

Fisher, Ronald A. (1956). *Statistical Methods and Scientific Inference.* Edinburgh: Oliver and Boyd.

Fitelson, Branden (2006). Logical Foundations of Evidential Support. *Philosophy of Science* 73, pp. 500–12.

Fitelson, Branden (2007). Likelihoodism, Bayesianism, and Relational Confirmation. *Synthese* 156, pp. 473–89.

Fitelson, Branden (2008). A Decision Procedure for Probability Calculus with Applications. *The Review of Symbolic Logic* 1, pp. 111–25.

Fitelson, Branden (2012). Evidence of Evidence Is Not (Necessarily) Evidence. *Analysis* 72, pp. 85–8.

Fitelson, Branden (2015). The Strongest Possible Lewisian Triviality Result. *Thought* 4, pp. 69–74.

Fitelson, Branden and Alan Hájek (2014). Declarations of Independence. *Synthese* 194, pp. 3979–95.

Fitelson, Branden and James Hawthorne (2010a). How Bayesian Confirmation Theory Handles the Paradox of the Ravens. *The Place of Probability in Science.* Ed. by Ellery Eells and J. Fetzer. *Boston Studies in the Philosophy of Science* 284, pp. 247–75.

Fitelson, Branden and James Hawthorne (2010b). The Wason Task(s) and the Paradox of Confirmation. *Philosophical Perspectives* 24, pp. 207–41.

Fitelson, Branden and David McCarthy (ms). Accuracy and Comparative Likelihood. Unpublished manuscript.

Fitelson, Branden and Daniel Osherson (2015). Remarks on "Random Sequences". *Australasian Journal of Logic* 12, pp. 11–16.

Foley, Richard (1987). *The Theory of Epistemic Rationality.* Cambridge, MA: Harvard University Press.

Foley, Richard (1993). *Working without a Net.* Oxford: Oxford University Press.

Foley, Richard (2009). Beliefs, Degrees of Belief, and the Lockean Thesis. In: *Degrees of Belief.* Ed. by Franz Huber and Christoph Schmidt-Petri. Vol. 342. Synthese Library. Springer, pp. 37–48.

Frigerio, Roberta et al. (2005). Education and Occupations Preceding Parkinson's Disease. *Neurology* 65, pp. 1575–83.

Gaifman, H. and M. Snir (1982). Probabilities over Rich Languages. *Journal of Symbolic Logic* 47, pp. 495–548.

Gaifman, Haim (2004). Reasoning with Limited Resources and Assigning Probabilities to Arithmetical Statements. *Synthese* 140, pp. 97–119.

Galavotti, Maria Carla (2005). *Philosophical Introduction to Probability.* CSLI Lecture Notes 167. Stanford, CA: CSLI Publications.

Gallo, Valentina et al. (2018). Exploring Causality of the Association between Smoking and Parkinson's Disease. *International Journal of Epidemiology* 48, pp. 912–25.

Gandenberger, Greg (2014). *Titelbaum's Counterexample to the Law of Likelihood.* Blog post available at URL: https://gandenberger.org/2014/05/26/titelbaum-counterexample/.

Gandenberger, Greg (2015). A New Proof of the Likelihood Principle. *British Journal for the Philosophy of Science* 66, pp. 475–503.

Gandenberger, Greg (ms). New Responses to Three Purported Counterexamples to the Likelihood Principle. Unpublished manuscript.

Garber, Daniel (1983). Old Evidence and Logical Omniscience in Bayesian Confirmation Theory. In: *Testing Scientific Theories*. Ed. by John Earman. Vol. 10. Minnesota Studies in the Philosophy of Science. Minneapolis: University of Minnesota Press, pp. 99–132.

Gibbard, A. and W. Harper (1978/1981). Counterfactuals and Two Kinds of Expected Utility. In: *Ifs: Conditionals, Belief, Decision, Chance, and Time*. Ed. by W. Harper, Robert C. Stalnaker, and G. Pearce. Dordrecht: Reidel, pp. 153–90.

Gillies, Donald (2000). Varieties of Propensity. *British Journal for the Philosophy of Science* 51, pp. 807–35.

Glass, David H. and Mark McCartney (2015). A New Argument for the Likelihood Ratio Measure of Confirmation. *Acta Analytica* 30, pp. 59–65.

Glymour, Clark (1980). *Theory and Evidence*. Princeton, NJ: Princeton University Press.

Good, I.J. (1952). Rational Decisions. *Journal of the Royal Statistical Society, Series B* 14, pp. 107–14.

Good, I.J. (1962). Subjective Probability as the Measure of a Non-measurable Set. In: *Logic, Methodology, and the Philosophy of Science*. Ed. by P. Suppes, E. Nagel, and A. Tarski. Stanford, CA: Stanford University Press, pp. 319–29.

Good, I.J. (1967). The White Shoe Is a Red Herring. *British Journal for the Philosophy of Science* 17, p. 322.

Good, I.J. (1968). The White Shoe qua Herring Is Pink. *British Journal for the Philosophy of Science* 19, pp. 156–7.

Good, I.J. (1971). Letter to the Editor. *The American Statistician* 25, pp. 62–3.

Goodman, Nelson (1946). A Query on Confirmation. *The Journal of Philosophy* 43, pp. 383–5.

Goodman, Nelson (1955). *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press.

Greaves, Hilary (2013). Epistemic Decision Theory. *Mind* 122, pp. 915–52.

Greaves, Hilary and David Wallace (2006). Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility. *Mind* 115, pp. 607–32.

Hacking, Ian (1965). *The Logic of Statistical Inference*. Cambridge: Cambridge University Press.

Hacking, Ian (1967). Slightly More Realistic Personal Probability. *Philosophy of Science* 34, pp. 311–25.

Hacking, Ian (1971). The Leibniz-Carnap Program for Inductive Logic. *The Journal of Philosophy* 68, pp. 597–610.

Hacking, Ian (2001). *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press.

Hájek, Alan (1996). 'Mises Redux'—Redux: Fifteen Arguments against Finite Frequentism. *Erkenntnis* 45, pp. 209–27.

Hájek, Alan (2003). What Conditional Probability Could Not Be. *Synthese* 137, pp. 273–323.

Hájek, Alan (2009a). Arguments for—or against—Probabilism? In: *Degrees of Belief*. Ed. by Franz Huber and Christoph Schmidt-Petri. Vol. 342. Synthese Library. Springer, pp. 229–51.

Hájek, Alan (2009b). Fifteen Arguments against Hypothetical Frequentism. *Erkenntnis* 70, pp. 211–35.

Hájek, Alan (2011a). Conditional Probability. In: *Philosophy of Statistics*. Ed. by Prasanta S. Bandyopadhyay and Malcolm R. Forster. Vol. 7. Handbook of the Philosophy of Science. Amsterdam: Elsevier, pp. 99–136.

Hájek, Alan (2011b). Triviality Pursuit. *Topoi* 30, pp. 3–15.

Hájek, Alan (2019). Interpretations of Probability. In: *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta. Fall 2019. URL: http://plato.stanford.edu/archives/fall2019/entries/probability-interpret/.

Hájek, Alan and James M. Joyce (2008). Confirmation. In: *The Routledge Companion to Philosophy of Science.* Ed. by Stathis Psillos and Martin Curd. New York: Routledge, pp. 115–28.

Hall, Ned (2004). Two Mistakes about Credence and Chance. *Australasian Journal of Philosophy* 82, pp. 93–111.

Halpern, Joseph Y. (1999). Cox's Theorem Revisited. *Journal of Artificial Intelligence Research* 11, pp. 429–35.

Halpern, Joseph Y. (2003). *Reasoning about Uncertainty.* Cambridge, MA: MIT Press.

Halpern, Joseph Y. (2004). Sleeping Beauty Reconsidered: Conditioning and Reflection in Asynchronous Systems. *Proceedings of the Twentieth Conference on Uncertainty in AI,* pp. 226–34.

Halpern, Joseph Y. (2005). Sleeping Beauty Reconsidered: Conditioning and Reflection in Asynchronous Systems. In: *Oxford Studies in Epistemology.* Ed. by Tamar Szabó Gendler and John Hawthorne. Vol. 1. Oxford: Oxford University Press, pp. 111–42.

Harman, Gilbert (1986). *Change in View.* Boston: The MIT Press.

Harrison-Trainor, Matthew, Wesley H. Holliday, and Thomas F. Icard III (2016). A Note on Cancellation Axioms for Comparative Probability. *Theory and Decision* 80, pp. 159–66.

Hart, Casey and Michael G. Titelbaum (2015). Intuitive Dilation? *Thought* 4, pp. 252–62.

Haverkamp, Nick and Moritz Schulz (2012). A Note on Comparative Probability. *Erkenntnis* 76, pp. 395–402.

Hawthorne, James (2014). Inductive Logic. In: *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta. Winter 2014. URL: http://plato.stanford.edu/archives/win2014/entries/logic-inductive/.

Hawthorne, James (2016). A Logic of Comparative Support: Qualitative Conditional Probability Relations Representable by Popper Functions. In: *The Oxford Handbook of Probability and Philosophy.* Ed. by Alan Hájek and Christopher R. Hitchcock. Oxford: Oxford University Press, pp. 277–95.

Hawthorne, James and Branden Fitelson (2004). Re-solving Irrelevant Conjunction with Probabilistic Independence. *Philosophy of Science* 71, pp. 505–14.

Hedden, Brian (2013). Incoherence without Exploitability. *Noûs* 47, pp. 482–95.

Hempel, Carl G. (1945a). Studies in the Logic of Confirmation (I). *Mind* 54, pp. 1–26.

Hempel, Carl G. (1945b). Studies in the Logic of Confirmation (II). *Mind* 54, pp. 97–121.

Hesse, Mary (1963). *Models and Analogies in Science.* London: Sheed & Ward.

Heukelom, Floris (2015). A History of the Allais Paradox. *The British Journal for the History of Science* 48, pp. 147–69.

Hintikka, Jaakko (1975). Impossible Possible Worlds Vindicated. *Journal of Philosophical Logic* 4, pp. 475–84.

Hitchcock, Christopher R. (2004). Beauty and the Bets. *Synthese* 139, pp. 405–20.

Hitchcock, Christopher R. (2021). Probabilistic Causation. In: *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta. Spring 2021. URL: https://plato.stanford.edu/archives/spr2021/entries/causation-probabilistic/.

Holton, Richard (2014). Intention as a Model for Belief. In: *Rational and Social Agency: The Philosophy of Michael Bratman.* Ed. by Manuel Vargas and Gideon Yaffe. Oxford: Oxford University Press, pp. 12–37.

Hooker, C.A. (1968). Goodman, 'Grue' and Hempel. *Philosophy of Science* 35, pp. 232–47.

Hosiasson-Lindenbaum, Janina (1940). On Confirmation. *Journal of Symbolic Logic* 5, pp. 133–48.

Howson, Colin (1992). Dutch Book Arguments and Consistency. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 2, pp. 161–8.

Howson, Colin (2014). Finite Additivity, Another Lottery Paradox and Conditionalisation. *Synthese* 191, pp. 989–1012.

Howson, Colin and Peter Urbach (2006). *Scientific Reasoning: The Bayesian Approach*. 3rd edition. Chicago: Open Court.

Huber, Franz (2016). Formal Representations of Belief. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2016. URL: https://plato.stanford.edu/archives/spr2016/entries/formal-belief/.

Hume, David (1739–40/1978). *A Treatise of Human Nature*. Ed. by L.A. Selby-Bigge and Peter H. Nidditch. 2nd edition. Oxford: Oxford University Press.

Humphreys, Paul (1985). Why Propensities Cannot Be Probabilities. *Philosophical Review* 94, pp. 557–70.

Icard III, Thomas F. (2016). Pragmatic Considerations on Comparative Probability. *Philosophy of Science* 83, pp. 348–70.

Jackson, Elizabeth G. (2020). The Relationship between Belief and Credence. *Philosophy Compass* 15, pp. 1–13.

Jaynes, E.T. (1957a). Information Theory and Statistical Mechanics I. *Physical Review* 106, pp. 62–30.

Jaynes, E.T. (1957b). Information Theory and Statistical Mechanics II. *Physical Review* 108, pp. 171–90.

Jaynes, E.T. (1968). Prior Probabilities. *IEEE Transactions on Systems Science and Cybernetics* SEC-4, pp. 227–41.

Jeffrey, Richard C. (1965). *The Logic of Decision*. 1st edition. McGraw-Hill Series in Probability and Statistics. New York: McGraw-Hill.

Jeffrey, Richard C. (1970). Dracula Meets Wolfman: Acceptance vs. Partial Belief. In: *Induction, Acceptance, and Rational Belief*. Ed. by M. Swain. Dordrecht: Reidel, pp. 157–85.

Jeffrey, Richard C. (1983). *The Logic of Decision*. 2nd edition. Chicago: University of Chicago Press.

Jeffrey, Richard C. (1987). Indefinite Probability Judgment: A Reply to Levi. *Philosophy of Science* 54, pp. 586–91.

Jeffrey, Richard C. (1993). Causality and the Logic of Decision. *Philosophical Topics* 21, pp. 139–51.

Jeffrey, Richard C. (2004). *Subjective Probability: The Real Thing*. Cambridge: Cambridge University Press.

Johnson, W.E. (1932). Probability: The Deductive and Inductive Problems. *Mind* 41, pp. 409–23.

Joyce, James M. (1998). A Nonpragmatic Vindication of Probabilism. *Philosophy of Science* 65, pp. 575–603.

Joyce, James M. (1999). *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.

Joyce, James M. (2005). How Probabilities Reflect Evidence. *Philosophical Perspectives* 19, pp. 153–78.

Joyce, James M. (2009). Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief. In: *Degrees of Belief*. Ed. by Franz Huber and Christoph Schmidt-Petri. Vol. 342. Synthese Library. Springer, pp. 263–97.

Joyce, James M. (2010). A Defense of Imprecise Credences in Inference and Decision Making. *Philosophical Perspectives* 24, pp. 281–323.

Kahneman, Daniel and Amos Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* XLVII, pp. 263–91.

Kaplan, Mark (1996). *Decision Theory as Philosophy.* Cambridge: Cambridge University Press.

Kelley, Mikayla (ms). On Accuracy and Coherence with Infinite Opinion Sets. Unpublished manuscript.

Kemeny, John G. (1955). Fair Bets and Inductive Probabilities. *The Journal of Symbolic Logic* 20, pp. 263–73.

Kemeny, John G. and Paul Oppenheim (1952). Degree of Factual Support. *Philosophy of Science* 19, pp. 307–24.

Keynes, John Maynard (1921). *Treatise on Probability.* London: Macmillan and Co., Limited.

Keynes, John Maynard (1923). *A Tract on Monetary Reform.* Macmillan and Co., Limited.

Keynes, John Maynard (1937). The General Theory of Employment. *The Quarterly Journal of Economics* 51, pp. 209–23.

Kierland, Brian and Bradley Monton (2005). Minimizing Inaccuracy for Self-locating Beliefs. *Philosophy and Phenomenological Research* 70, pp. 384–95.

Kim, Jaegwon (1988). What Is "Naturalized Epistemology"? *Philosophical Perspectives* 2, pp. 381–405.

Kim, Namjoong (2009). Sleeping Beauty and Shifted Jeffrey Conditionalization. *Synthese* 168, pp. 295–312.

Knight, Frank Hyneman (1921). *Risk, Uncertainty and Profit.* Boston and New York: Houghton Mifflin.

Kolmogorov, A.N. (1933/1950). *Foundations of the Theory of Probability.* Translation edited by Nathan Morrison. New York: Chelsea Publishing Company.

Konek, Jason and Benjamin A. Levinstein (2019). The Foundations of Epistemic Decision Theory. *Mind* 128, pp. 69–107.

Kornblith, Hilary (1993). Epistemic Normativity. *Synthese* 94, pp. 357–76.

Kraft, Charles H., John W. Pratt, and A. Seidenberg (1959). Intuitive Probability on Finite Sets. *The Annals of Mathematical Statistics* 30, pp. 408–19.

Krantz, D.H. et al. (1971). *Foundations of Measurement, Vol 1: Additive and Polynomial Representations*. Cambridge: Academic Press.

Kuhn, Thomas S. (1957). *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought.* New York: MJF Books.

Kyburg Jr, Henry E. (1961). *Probability and the Logic of Rational Belief.* Middletown: Wesleyan University Press.

Kyburg Jr, Henry E. (1970). Conjunctivitis. In: *Induction, Acceptance, and Rational Belief.* Ed. by M. Swain. Boston: Reidel, pp. 55–82.

Laddaga, R. (1977). Lehrer and the Consensus Proposal. *Synthese* 36, pp. 473–7.

Lance, Mark Norris (1995). Subjective Probability and Acceptance. *Philosophical Studies* 77, pp. 147–79.

Lange, Alexandra (2019). Can Data Be Human? The Work of Giorgia Lupi. *The New Yorker.* Published May 25, 2019.

Lange, Marc (2000). Is Jeffrey Conditionalization Defective by Virtue of Being Non-commutative? Remarks on the Sameness of Sensory Experience. *Synthese* 123, pp. 393–403.

Laplace, Pierre-Simon (1814/1995). *Philosophical Essay on Probabilities.* Translated from the French by Andrew Dale. New York: Springer.

Lehman, R. Sherman (1955). On Confirmation and Rational Betting. *Journal of Symbolic Logic* 20, pp. 251–62.

Lehrer, K. and Carl Wagner (1983). Probability Amalgamation and the Independence Issue: A Reply to Laddaga. *Synthese* 55, pp. 339–46.

Leitgeb, Hannes and Richard Pettigrew (2010a). An Objective Justification of Bayesianism I: Measuring Inaccuracy. *Philosophy of Science* 77, pp. 201–35.

Leitgeb, Hannes and Richard Pettigrew (2010b). An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy. *Philosophy of Science* 77, pp. 236–72.

Lele, Subhash R. (2004). Evidence Functions and the Optimality of the Law of Likelihood. In: *The Nature of Scientific Evidence: Statistical, Philosophical, and Empirical Considerations*. Chicago: University of Chicago Press, pp. 191–216.

Levi, Isaac (1974). On Indeterminate Probabilities. *The Journal of Philosophy* 71, pp. 391–418.

Levi, Isaac (1980). *The Enterprise of Knowledge*. Boston: The MIT Press.

Levi, Isaac (1987). The Demons of Decision. *The Monist* 70, pp. 193–211.

Lewis, C.I. (1946). *An Analysis of Knowledge and Valuation*. La Salle, IL: Open Court.

Lewis, David (1971). Immodest Inductive Methods. *Philosophy of Science* 38, pp. 54–63.

Lewis, David (1976). Probabilities of Conditionals and Conditional Probabilities. *The Philosophical Review* 85, pp. 297–315.

Lewis, David (1979). Atittudes *de dicto* and *de se*. *The Philosophical Review* 88, pp. 513–43.

Lewis, David (1980). A Subjectivist's Guide to Objective Chance. In: *Studies in Inductive Logic and Probability*. Ed. by Richard C. Jeffrey. Vol. 2. Berkeley: University of California Press, pp. 263–94.

Lewis, David (1981a). Causal Decision Theory. *Australasian Journal of Philosophy* 59, pp. 5–30.

Lewis, David (1981b). 'Why Ain'cha Rich?' *Noûs* 15, pp. 377–80.

Lewis, David (1994). Humean Supervenience Debugged. *Mind* 103, pp. 473–90.

Lewis, David (1996). Desire as Belief II. *Mind* 105, pp. 303–13.

Lewis, David (2001). Sleeping Beauty: Reply to Elga. *Analysis* 61, pp. 171–6.

Lichtenstein, S., B. Fischoff, and L. Phillips (1982). Calibration of Probabilities: The State of the Art to 1980. In: *Judgment under Uncertainty: Heuristics and Biases*. Ed. by Daniel Kahneman, P. Slovic, and Amos Tversky. Cambridge: Cambridge University Press, pp. 306–34.

Lindley, Dennis V. (1982). Scoring Rules and the Inevitability of Probability. *International Statistical Review* 50, pp. 1–26.

Lindley, Dennis V. (1985). *Making Decisions*. 2nd edition. London: Wiley.

Liu, Liping and Ronald R. Yager (2008). Classic Works of the Dempster-Shafer Theory of Belief Functions: An Introduction. In: *Classic Works of the Dempster-Shafer Theory of Belief Functions*. Ed. by Ronald R. Yager and Liping Liu. Vol. 219. Studies in Fuzziness and Soft Computing. Berlin: Springer, pp. 1–34.

Locke, John (1689/1975). *An Essay Concerning Human Understanding*. Ed. by Peter H. Nidditch. Oxford: Oxford University Press.

MacFarlane, John (2005). Making Sense of Relative Truth. *Proceedings of the Aristotelian Society* 105, pp. 321–39.

Maher, Patrick (1993). *Betting on Theories*. Cambridge Studies in Probability, Induction, and Decision Theory. Cambridge: Cambridge University Press.

Maher, Patrick (1996). Subjective and Objective Confirmation. *Philosophy of Science* 63, pp. 149–74.

Maher, Patrick (2002). Joyce's Argument for Probabilism. *Philosophy of Science* 96, pp. 73–81.

Maher, Patrick (2010). Explication of Inductive Probability. *Journal of Philosophical Logic* 39, pp. 593–616.

Makinson, David C. (1965). The Paradox of the Preface. *Analysis* 25, pp. 205–7.

Makinson, David C. (2011). Conditional Probability in the Light of Qualitative Belief Change. *Journal of Philosophical Logic* 40, pp. 121–53.

Mayo, Deborah (2018). *Statistical Inference as Severe Testing: How to Get beyond the Statistics Wars*. Cambridge: Cambridge University Press.

Mazurkiewicz, Stefan (1932). Zur Axiomatik der Wahrscheinlichkeitsrechnung. *Comptes rendues des séances de la Société des Sciences et des Lettres de Varsovie* 25, pp. 1–4.

Meacham, Christopher J. G. (2008). Sleeping Beauty and the Dynamics of De Se Beliefs. *Philosophical Studies* 138, pp. 245–70.

Meacham, Christopher J. G. (2010a). Unravelling the Tangled Web: Continuity, Internalism, Uniqueness, and Self-locating Belief. In: *Oxford Studies in Epistemology*. Ed. by Tamar Szabó Gendler and John Hawthorne. Vol. 3. Oxford: Oxford University Press, pp. 86–125.

Meacham, Christopher J.G. (2010b). Two Mistakes Regarding the Principal Principle. *British Journal for the Philosophy of Science* 61, pp. 407–31.

Meacham, Christopher J. G. (2016). Ur-Priors, Conditionalization, and Ur-Prior Conditionalization. *Ergo* 3, pp. 444–92.

Meacham, Christopher J.G. and Jonathan Weisberg (2011). Representation Theorems and the Foundations of Decision Theory. *Australasian Journal of Philosophy* 89, pp. 641–63.

Mellor, D.H. (2013). Review of *Probability in the Philosophy of Religion*. *Analysis* 73, pp. 548–54.

Moore, David S., George P. McCabe, and Bruce A. Craig (2009). *Introduction to the Practice of Statistics*. 6th edition. New York: W.H. Freeman and Company.

Moore, G.E. (1939). Proof of an External World. *Proceedings of the British Academy* 25, pp. 273–300.

Moss, Sarah (2012). Updating as Communication. *Philosophy and Phenomenological Research* 85, pp. 225–48.

Moss, Sarah (2015). Credal Dilemmas. *Noûs* 49, pp. 665–83.

Moss, Sarah (2018). *Probabilistic Knowledge*. Oxford: Oxford University Press.

Murphy, Allan H. (1973). A New Vector Partition of the Probability Score. *Journal of Applied Meteorology* 12, pp. 595–600.

Murphy, Allan H. and Robert L. Winkler (1977). Reliability of Subjective Probability Forecasts of Precipitation and Temperature. *Journal of the Royal Statistical Society, Series C* 26, pp. 41–7.

Neal, Radford M. (2006). *Puzzles of Anthropic Reasoning Resolved Using Full Non-indexical Conditioning*. Tech. rep. 0607. Department of Statistics, University of Toronto.

Neyman, J. and Egon Pearson (1967). *Joint Statistical Papers*. Cambridge: Cambridge University Press.

Nicod, Jean (1930). *Foundations of Geometry and Induction*. Translated by Philip Wiener. New York: Harcourt, Brace and Company.

Nozick, Robert (1969). Newcomb's Problem and Two Principles of Choice. In: *Essays in Honor of Carl G. Hempel*. Synthese Library. Dordrecht: Reidel, pp. 114–15.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science* 349. DOI: 10.1126/science.aac4716.

Papineau, David (2012). *Philosophical Devices: Proofs, Probabilities, Possibilities, and Sets*. Oxford: Oxford University Press.

Paris, J.B. (1994). *The Uncertain Reasoner's Companion*. Cambridge: Cambridge University Press.

Pascal, Blaise (1670/1910). *Pensées*. Translated by W.F. Trotter. London: Dent.

Pearson, K., A. Lee, and L. Bramley-Moore (1899). Genetic (Reproductive) Selection: Inheritance of Fertility in Man. *Philosophical Transactions of the Royal Society A* 73, pp. 534–39.

Pedersen, Arthur Paul and Gregory Wheeler (2014). Demystifying Dilation. *Erkenntnis* 79, pp. 1305–42.

Peirce, Charles Sanders (1910/1932). Notes on the Doctrine of Chances. In: *Collected Papers of Charles Sanders Peirce*. Ed. by Charles Hartshorne and Paul Weiss. Cambridge, MA: Harvard University Press, pp. 404–14.

Peterson, Martin (2009). *An Introduction to Decision Theory*. Cambridge Introductions to Philosophy. Cambridge: Cambridge University Press.

Pettigrew, Richard (2013a). A New Epistemic Utility Argument for the Principal Principle. *Episteme* 10, pp. 19–35.

Pettigrew, Richard (2013b). Epistemic Utility and Norms for Credences. *Philosophy Compass* 8, pp. 897–908.

Pettigrew, Richard (2014). Accuracy, Risk, and the Principle of Indifference. *Philosophy and Phenomenological Research* 92(1), pp. 35–59.

Pettigrew, Richard (2016). *Accuracy and the Laws of Credence*. Oxford: Oxford University Press.

Pettigrew, Richard (2021). On the Expected Utility Objection to the Dutch Book Argument for Probabilism. *Noûs* 55, pp. 23–38.

Pettigrew, Richard and Michael G. Titelbaum (2014). Deference Done Right. *Philosophers' Imprint* 14, pp. 1–19.

Pettigrew, Richard and Jonathan Weisberg, eds (2019). *The Open Handbook of Formal Epistemology*. Published open access online by PhilPapers. URL: https://philpapers.org/rec/PETTOH-2.

Piccione, Michele and Ariel Rubinstein (1997). On the Interpretation of Decision Problems with Imperfect Recall. *Games and Economic Behavior* 20, pp. 3–24.

Pollock, John L. (2001). Defeasible Reasoning with Variable Degrees of Justification. *Artificial Intelligence* 133, pp. 233–82.

Popper, Karl R. (1935/1959). *The Logic of Scientific Discovery*. London: Hutchinson & Co.

Popper, Karl R. (1938). A Set of Independent Axioms for Probability. *Mind* 47, pp. 275–9.

Popper, Karl R. (1954). Degree of Confirmation. *British Journal for the Philosophy of Science* 5, pp. 143–9.

Popper, Karl R. (1955). Two Autonomous Axiom Systems for the Calculus of Probabilities. *British Journal for the Philosophy of Science* 6, pp. 51–7.

Popper, Karl R. (1957). The Propensity Interpretation of the Calculus of Probability and the Quantum Theory. *The Colston Papers* 9. Ed. by S. Körner, pp. 65–70.

Predd, J. et al. (2009). Probabilistic Coherence and Proper Scoring Rules. *IEEE Transactions on Information Theory* 55, pp. 4786–92.

Pryor, James (2004). What's Wrong with Moore's Argument? *Philosophical Issues* 14, pp. 349–78.

Quinn, Warren S. (1990). The Puzzle of the Self-Torturer. *Philosophical Studies* 59, pp. 79–90.

Ramsey, Frank P. (1929/1990). General Propositions and Causality. In: *Philosophical Papers*. Ed. by D.H. Mellor. Cambridge: Cambridge University Press, pp. 145–63.

Ramsey, Frank P. (1931). Truth and Probability. In: *The Foundations of Mathematics and other Logic Essays*. Ed. by R.B. Braithwaite. New York: Harcourt, Brace and Company, pp. 156–98.

Rees, Martin (2000). *Just Six Numbers: The Deep Forces that Shape the Universe*. New York: Basic Books.

Reichenbach, Hans (1935/1949). *The Theory of Probability.* English expanded version of the German original. Berkeley: University of California Press.

Reichenbach, Hans (1938). *Experience and Prediction.* Chicago: University of Chicago Press.

Reichenbach, Hans (1956). The Principle of Common Cause. In: *The Direction of Time.* Berkeley: University of California Press, pp. 157–60.

Reiss, Julian and Jan Sprenger (2017). Scientific Objectivity. In: *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta. Winter 2017. Metaphysics Research Lab, Stanford University. URL: https://plato.stanford.edu/archives/win2017/entries/scientific-objectivity/.

Renyi, Alfred (1970). *Foundations of Probability.* San Francisco: Holden-Day.

Resnik, Michael D. (1987). *Choices: An Introduction to Decision Theory.* Minneapolis: University of Minnesota Press.

Roche, William (2014). Evidence of Evidence Is Evidence under Screening-off. *Episteme* 11, pp. 119–24.

Roeper, P. and H. Leblanc (1999). *Probability Theory and Probability Logic.* Toronto: University of Toronto Press.

Rosenkrantz, Roger (1981). *Foundations and Applications of Inductive Probability.* Atascadero, CA: Ridgeview Press.

Rosenthal, Jeffrey S. (2009). A Mathematical Analysis of the Sleeping Beauty Problem. *The Mathematical Intelligencer* 31, pp. 32–37.

Royall, Richard M. (1997). *Statistical Evidence: A Likelihood Paradigm.* New York: Chapman & Hall/CRC.

Salmon, Wesley (1966). *The Foundations of Scientific Inference.* Pittsburgh: University of Pittsburgh Press.

Salmon, Wesley (1975). Confirmation and Relevance. In: *Induction, Probability, and Confirmation.* Ed. by Grover Maxwell and Robert M. Jr. Anderson. Vol. VI. Minnesota Studies in the Philosophy of Science. Minneapolis: University of Minnesota Press, pp. 3–36.

Savage, Leonard J. (1954). *The Foundations of Statistics.* New York: Wiley.

Savage, Leonard J. (1967). Difficulties in the Theory of Personal Probability. *Philosophy of Science* 34, pp. 305–10.

Schervish, M.J., T. Seidenfeld, and J.B. Kadane (2004). Stopping to Reflect. *Journal of Philosophy* 101, pp. 315–22.

Schervish, M.J., T. Seidenfeld, and J.B. Kadane (2009). Proper Scoring Rules, Dominated Forecasts, and Coherence. *Decision Analysis* 6, pp. 202–21.

Schick, Frederic (1986). Dutch Bookies and Money Pumps. *The Journal of Philosophy* 83, pp. 112–19.

Schoenfield, Miriam (2017). The Accuracy and Rationality of Imprecise Credences. *Noûs* 51, pp. 667–85.

Schwarz, Wolfgang (2010). *Lewis on Updating and Self-location.* Blog post available at URL: https://www.umsu.de/wo/2010/563.

Schwarz, Wolfgang (2018). Subjunctive Conditional Probability. *Journal of Philosophical Logic* 47, pp. 47–66.

Scott, Dana (1964). Measurement Structures and Linear Inequalities. *Journal of Mathematical Psychology* 1, pp. 233–47.

Seelig, C. (1956). *Albert Einstein: A Documentary Biography.* London: Staples Press.

Seidenfeld, Teddy (1986). Entropy and Uncertainty. *Philosophy of Science* 53, pp. 467–91.

Seidenfeld, Teddy, M.J. Schervish, and J.B. Kadane (2017). Non-Conglomerability for Countably Additive Measures That Are Not κ-Additive. *The Review of Symbolic Logic* 10, pp. 284–300.

Seidenfeld, Teddy and Larry Wasserman (1993). Dilation for Sets of Probabilities. *The Annals of Statistics* 21, pp. 1139–54.

Selvin, Steve (1975). A Problem in Probability. *The American Statistician* 29. Published among the Letters to the Editor, p. 67.

Shafer, Glenn (1976). *A Mathematical Theory of Evidence.* Princeton, NJ: Princeton University Press.

Shafer, Glenn (1981). Constructive Probability. *Synthese* 48, pp. 1–60.

Shapiro, Amram, Louise Firth Campbell, and Rosalind Wright (2014). *The Book of Odds.* New York: Harper Collins.

Shimony, Abner (1955). Coherence and the Axioms of Confirmation. *Journal of Symbolic Logic* 20, pp. 1–28.

Shimony, Abner (1988). An Adamite Derivation of the Calculus of Probability. In: *Probability and Causality.* Ed. by J.H. Fetzer. Dordrecht: Reidel, pp. 151–61.

Shogenji, Tomoji (2003). A Condition for Transitivity in Probabilistic Support. *British Journal for the Philosophy of Science* 54, pp. 613–16.

Shogenji, Tomoji (2012). The Degree of Epistemic Justification and the Conjunction Fallacy. *Synthese* 184, pp. 29–48.

Shope, R.K. (1978). The Conditional Fallacy in Contemporary Philosophy. *Journal of Philosophy* 75, pp. 397–413.

Shortliffe, E. and B. Buchanan (1975). A Model of Inexact Reasoning in Medicine. *Mathematical Biosciences* 23, pp. 351–79.

Simpson, E.H. (1951). The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society, Series B* 13, pp. 238–41.

Skipper, Mattias and Jens Christian Bjerring (2022). Bayesianism for Nonideal Agents. *Erkenntnis* 87, pp. 93–115.

Skyrms, Brian (1980a). *Causal Necessity: A Pragmatic Investigation of the Necessity of Laws.* New Haven CT: Yale University Press.

Skyrms, Brian (1980b). Higher Order Degrees of Belief. In: *Prospects for Pragmatism.* Ed. by D.H. Mellor. Cambridge: Cambridge University Press, pp. 109–37.

Skyrms, Brian (1983). Three Ways to Give a Probability Function a Memory. In: *Testing Scientific Theories.* Ed. by John Earman. Vol. 10. Minnesota Studies in the Philosophy of Science. Minneapolis: University of Minnesota Press, pp. 157–61.

Skyrms, Brian (1987a). Coherence. In: *Scientific Inquiry in Philosophical Perspective.* Ed. by N. Rescher. Pittsburgh: University of Pittsburgh Press, pp. 225–42.

Skyrms, Brian (1987b). Dynamic Coherence and Probability Kinematics. *Philosophy of Science* 54, pp. 1–20.

Skyrms, Brian (2000). *Choice & Chance: An Introduction to Inductive Logic.* 4th edition. Stamford, CT: Wadsworth.

Smith, C.A.B. (1961). Consistency in Statistical Inference and Decision. *Journal of the Royal Statistical Society, Series B* 23, pp. 1–25.

Smithies, Declan (2015). Ideal Rationality and Logical Omniscience. *Synthese* 192, pp. 2769–93.

Sober, Elliott (2008). *Evidence and Evolution.* Cambridge: Cambridge University Press.

Spohn, Wolfgang (2012). *The Laws of Belief: Ranking Theory and its Philosophical Applications.* Oxford: Oxford University Press.

Staffel, Julia (2019). *Unsettled Thoughts: Reasoning and Uncertainty in Epistemology.* Oxford: Oxford University Press.

Stalnaker, Robert C. (1972/1981). Letter to David Lewis. In: *Ifs: Conditionals, Belief, Decision, Chance, and Time.* Ed. by W. Harper, Robert C. Stalnaker, and G. Pearce. Dordrecht: Reidel, pp. 151–2.

Stalnaker, Robert C. (2008). *Our Knowledge of the Internal World*. Oxford: Oxford University Press.

Stalnaker, Robert C. (2011). Responses to Stoljar, Weatherson and Boghossian. *Philosophical Studies* 155, pp. 467–79.

Stefánsson, H. Orri (2017). What Is "Real" in Probabilism? *Australasian Journal of Philosophy* 95, pp. 573–87.

Stefánsson, H. Orri (2018). On the Ratio Challenge for Comparativism. *Australasian Journal of Philosophy* 96, pp. 380–90.

Stephenson, Todd A. (2000). *An Introduction to Bayesian Network Theory and Usage*. Tech. rep. 03. IDIAP.

Sturgeon, Scott (2010). Confidence and Coarse-grained Attitudes. In: *Oxford Studies in Epistemology*. Ed. by Tamar Szabó Gendler and John Hawthorne. Vol. 3. Oxford: Oxford University Press, pp. 126–49.

Suppes, Patrick (1955). The Role of Subjective Probability and Utility in Decision-making. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, pp. 61–73.

Suppes, Patrick (1969). *Studies in the Methodology and Foundations of Science*. Berlin: Springer.

Suppes, Patrick (1974). *Probabilistic Metaphysics*. Uppsala: University of Uppsala Press.

Suppes, Patrick (2002). *Representation and Invariance of Scientific Structures*. Stanford, CA: CSLI Publications.

Tal, Eyal and Juan Comesaña (2017). Is Evidence of Evidence Evidence? *Noûs* 51, pp. 95–112.

Talbott, William J. (1991). Two Principles of Bayesian Epistemology. *Philosophical Studies* 62, pp. 135–50.

Talbott, William J. (2005). Review of "Putting Logic in its Place: Formal Constraints on Rational Belief". *Notre Dame Philosophical Reviews*. URL: http://ndpr.nd.edu/.

Talbott, William J. (2016). Bayesian Epistemology. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2016. Metaphysics Research Lab, Stanford University. URL: https://plato.stanford.edu/archives/win2016/entries/epistemology-bayesian/.

Teller, Paul (1973). Conditionalization and Observation. *Synthese* 26, pp. 218–58.

Tennant, Neil (2017). Logicism and Neologicism. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2017. Metaphysics Research Lab, Stanford University. URL: https://plato.stanford.edu/archives/win2017/entries/logicism/.

Tentori, Katya, Vincenzo Crupi, and Selena Russo (2013). On the Determinants of the Conjunction Fallacy: Probability versus Inductive Confirmation. *Journal of Experimental Psychology: General* 142, pp. 235–55.

Titelbaum, Michael G. (2008). The Relevance of Self-locating Beliefs. *Philosophical Review* 117, pp. 555–605.

Titelbaum, Michael G. (2010). Not Enough There There: Evidence, Reasons, and Language Independence. *Philosophical Perspectives* 24, pp. 477–528.

Titelbaum, Michael G. (2013a). *Quitting Certainties: A Bayesian Framework Modeling Degrees of Belief*. Oxford: Oxford University Press.

Titelbaum, Michael G. (2013b). Ten Reasons to Care about the Sleeping Beauty Problem. *Philosophy Compass* 8, pp. 1003–17.

Titelbaum, Michael G. (2015a). Continuing On. *Canadian Journal of Philosophy* 45, pp. 670–91.

Titelbaum, Michael G. (2015b). Rationality's Fixed Point (Or: In Defense of Right Reason). In: *Oxford Studies in Epistemology*. Ed. by Tamar Szabó Gendler and John Hawthorne. Vol. 5. Oxford University Press, pp. 253–94.

Titelbaum, Michael G. (2016). Self-locating Credences. In: *The Oxford Handbook of Probability and Philosophy.* Ed. by Alan Hájek and Christopher R. Hitchcock. Oxford: Oxford University Press, pp. 666–80.

Titelbaum, Michael G. (2019). Precise Credences. In: *The Open Handbook of Formal Epistemology.* Ed. by Richard Pettigrew and Jonathan Weisberg. The PhilPapers Foundation, pp. 1–56.

Tversky, Amos and Daniel Kahneman (1974). Judgment under Uncertainty: Heuristics and Biases. *Science* 185, pp. 1124–31.

Tversky, Amos and Daniel Kahneman (1983). Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment. *Psychological Review* 90, pp. 293–315.

Tversky, Amos and Daniel Kahneman (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty* 5, pp. 297–323.

Tversky, Amos and Derek J. Koehler (1994). Support Theory: A Nonextensional Representation of Subjective Probability. *Psychological Review* 101, pp. 547–67.

van Enk, Steven J. (2015). Betting, Risk, and the Law of Likelihood. *Ergo* 2, pp. 105–21.

van Fraassen, Bas C. (1980). Rational Belief and Probability Kinematics. *Philosophy of Science* 47, pp. 165–87.

van Fraassen, Bas C. (1981). A Problem for Relative Information Minimizers. *British Journal for the Philosophy of Science* 32, pp. 375–79.

van Fraassen, Bas C. (1982). Rational Belief and the Common Cause Principle. In: *What? Where? When? Why?* Ed. by Robert McLaughlin. Dordrecht: Reidel, pp. 193–209.

van Fraassen, Bas C. (1983). Calibration: A Frequency Justification for Personal Probability. In: *Physics Philosophy and Psychoanalysis.* Ed. by R. Cohen and L. Laudan. Dordrecht: Reidel, pp. 295–319.

van Fraassen, Bas C. (1984). Belief and the Will. *The Journal of Philosophy* 81, pp. 235–56.

van Fraassen, Bas C. (1989). *Laws and Symmetry.* Oxford: Clarendon Press.

van Fraassen, Bas C. (1995). Belief and the Problem of Ulysses and the Sirens. *Philosophical Studies* 77, pp. 7–37.

van Fraassen, Bas C. (1999). Conditionalization: A New Argument For. *Topoi* 18, pp. 93–6.

van Fraassen, Bas C. (2005). Conditionalizing on Violated Bell's Inequalities. *Analysis* 65, pp. 27–32.

Van Horn, Kevin S. (2003). Constructing a Logic of Plausible Inference: A Guide to Cox's Theorem. *International Journal of Approximate Reasoning* 34, pp. 3–24.

Venn, John (1866). *The Logic of Chance.* London, Cambridge: Macmillan.

Vineberg, Susan (2011). Dutch Book Arguments. In: *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta. Summer 2011. URL: https://plato.stanford.edu/archives/su2011/entries/dutch-book/.

von Mises, Richard (1928/1957). *Probability, Statistics and Truth.* (English edition of the original German *Wahrscheinlichkeit, Statistik und Wahrheit.*) New York: Dover.

von Neumann, J. and O. Morgenstern (1947). *Theory of Games and Economic Behavior.* 2nd edition. Princeton, NJ: Princeton University Press.

Vranas, Peter B.M. (2004). Hempel's Raven Paradox: A Lacuna in the Standard Bayesian Solution. *British Journal for the Philosophy of Science* 55, pp. 545–60.

Wainer, Howard (2011). *Uneducated Guesses: Using Evidence to Uncover Misguided Education Policies.* Princeton, NJ: Princeton University Press.

Walley, Peter (1991). *Statistical Reasoning with Imprecise Probabilities.* London: Chapman and Hall.

Wasserstein, Ronald L. and Nicole A. Lazar (2016). The ASA's Statement on p-values: Context, Process, and Purpose. *The American Statistician.* DOI: 10.1080/00031305.2016.1154108.

Weatherson, Brian (2002). Keynes, Uncertainty and Interest Rates. *Cambridge Journal of Economics* 26, pp. 47–62.

Weatherson, Brian (2011). Stalnaker on Sleeping Beauty. *Philosophical Studies* 155, pp. 445–56.

Weatherson, Brian (2015). For Bayesians, Rational Modesty Requires Imprecision. *Ergo* 2, pp. 529–45.

Weatherson, Brian and Andy Egan (2011). Epistemic Modals and Epistemic Modality. In: *Epistemic Modality.* Ed. by Andy Egan and Brian Weatherson. Oxford: Oxford University Press, pp. 1–18.

Weintraub, Ruth (2001). The Lottery: A Paradox Regained and Resolved. *Synthese* 129, pp. 439–49.

Weirich, Paul (2012). Causal Decision Theory. In: *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta. Winter 2012. URL: https://plato.stanford.edu/archives/win2012/entries/decision-causal/.

Weisberg, Jonathan (2007). Conditionalization, Reflection, and Self-Knowledge. *Philosophical Studies* 135, pp. 179–97.

Weisberg, Jonathan (2009). Varieties of Bayesianism. In: *Handbook of the History of Logic.* Ed. by Dov. M Gabbya, Stephan Hartmann, and John Woods. Vol. 10: Inductive Logic. Oxford: Elsevier.

White, Roger (2005). Epistemic Permissiveness. *Philosophical Perspectives* 19, pp. 445–59.

White, Roger (2006). Problems for Dogmatism. *Philosophical Studies* 131, pp. 525–57.

White, Roger (2010). Evidential Symmetry and Mushy Credence. In: *Oxford Studies in Epistemology.* Ed. by Tamar Szabó Gendler and John Hawthorne. Vol. 3. Oxford: Oxford University Press, pp. 161–86.

Williams, J. Robert G. (ms). A Non-Pragmatic Dominance Argument for Conditionalization. Unpublished manuscript.

Williams, J. Robert G. (2016). Probability and Non-Classical Logic. In: *Oxford Handbook of Probability and Philosophy.* Ed. by Alan Hájek and Christopher R. Hitchcock. Oxford: Oxford University Press.

Williamson, Timothy (2000). *Knowledge and its Limits.* Oxford: Oxford University Press.

Williamson, Timothy (2007). How Probable Is an Infinite Sequence of Heads? *Analysis* 67, pp. 173–80.

Wittgenstein, Ludwig (1921/1961). *Tractatus Logico-Philosophicus.* Translated by D.F. Pears and B.F. McGuinness. London: Routledge.

Wong, S.M. et al. (1991). Axiomatization of Qualitative Belief Structure. *IEEE Transactions on Systems, Man and Cybernetics* 21, pp. 726–34.

Woolston, Chris (2015). Psychology Journal Bans *P* Values. *Nature* 519.9. URL: https://doi.org/10.1038/519009f.

Wroński, Leszek and Godziszewski, Michał Tomasz (2017). *The Stubborn Non-probabilist—'Negation Incoherence' and a New Way to Block the Dutch Book Argument.* International Workshop on Logic, Rationality and Interaction. Berlin, Heidelberg: Springer, pp. 256–67.

Yalcin, Seth (2012). A Counterexample to Modus Tollens. *Journal of Philosophical Logic* 41, pp. 1001–24.

Yule, G.U. (1903). Notes on the Theory of Association of Attributes in Statistics. *Biometrika* 2, pp. 121–34.

Zynda, Lyle (1995). Old Evidence and New Theories. *Philosophical Studies* 77, pp. 67–95.

Zynda, Lyle (2000). Representation Theorems and Realism about Degrees of Belief. *Philosophy of Science* 67, pp. 45–69.